

# Schools and Social Capital: Economic Segregation and Long-term Outcomes\*

Farah Mallah

September 8, 2024

## Abstract

How does k-12 school economic segregation impact lower-income students' college enrollment and employment prospects? Using data from Texas, I find that the typical (median) lower-income student goes through grades 5 to 12 having 1 in 20 upper-income classmates compared to 1 in 2 for the typical upper-income student. The low exposure to upper-income classmates is mainly driven by residential and school socioeconomic segregation, classroom assignment contributes little. I use within-school residual variation from school trends in the proportion of upper-income students to capture the impact of having a higher share of upper-income peers. I distinguish between two possible explanations for the relationship between exposure to upper-income classmates and long-term outcomes: spillover from peers with higher family income and spillover from higher-achieving peers. The slope of the relationship between the proportion of upper-income peers and college enrollment and wages implies that, in the absence of income segregation, lower-income students' 4-year college enrollment would increase by 0.44 percentage points and their young adult quarterly wages would increase by 1.9%, on average. The impact of exposure to upper-income peers appears to be independent of peer academic achievement.

---

\*Please do not circulate. This draft is for internal use only. I am thankful to my doctoral advisor Eric Taylor and my committee members Christopher Avery, Susan Dynarski and Lawrence Katz for their continued guidance, feedback, and comments throughout different stages of this project. I would also like to thank Thomas Kane, Martin West and Benjamin Goldman for reading versions of the paper and providing feedback and labor/public workshop attendees including Raj Chetty, Amanda Pallais, Edward Glaeser and David Cutler for their helpful comments and suggestions. This research is supported by research grants from the Russell Sage Foundation, Institute of Education Sciences (U.S. Department of Education), and Institute for Quantitative Social Science and the Stone Research Grant from Harvard Kennedy School's James M. and Cathleen D. Stone Program in Wealth Distribution, Inequality, and Social Policy. Any opinions expressed are mine alone and should not be construed as representing the opinions of the Russell Sage Foundation or any other organization. The conclusions of this research do not necessarily reflect the opinion or official position of the Texas Education Research Center, the Texas Education Agency, the Texas Higher Education Coordinating Board, the Texas Workforce Commission, or the State of Texas.

## 1 Introduction

An important goal of education is economic mobility, but the likelihood of a U.S. child earning more than her parents has declined in recent decades (Chetty et al., 2017). One factor long theorized to relate to economic mobility is social capital, defined as relationships with others that enable individuals to access resources (Coleman, 1988).<sup>1</sup> Most recently, Chetty et al. (2022) revealed that friendships with higher-income individuals (cross-income social capital) account for a substantial portion of the variation in economic mobility across neighborhoods.

However, it remains unclear how the way schools are organized may hinder/enable cross-income friendships and why cross-income friendships might matter for economic mobility. The literature on school socioeconomic segregation suggests that cross-income exposure is not common (Owens, Reardon and Jencks, 2016; Dalane and Marcotte, 2020; Kalogrides and Loeb, 2013). Having more upper-income classmates could impact a lower-income student's future prospects by granting her access to resources (e.g., teacher quality and school spending) and/or facilitating peer interactions and access to more valuable social networks. Understanding what shapes cross-income exposure and why cross-income exposure might matter for economic mobility has important policy implications. For example, if the relationship between exposure to upper-income peers and economic mobility operates exclusively through access to school resources, then policies that equalize access to school resources may improve economic mobility without addressing school economic segregation.

To illuminate how economic segregation in k-12 schools impacts lower-income students' college enrollment and employment prospects, I document the contribution of district, school and classroom economic segregation on lower-income students' exposure to upper-income peers and the impact of upper-income peers on college enrollment and employment. I do this in two steps. First, I present descriptive facts on lower-income students' proportion of upper-income classmates between grades 5 and 12 using data from Texas. I quantify the role of districts, schools and classrooms in accounting for differences in exposure to upper-income peers. I then capture how exposure to upper-income peers relates to friendship formation using the Add Health data and Chetty et al. (2022) school-level cross-income friendship measures. The second part of the paper estimates the relationship between exposure to upper-income students and students' college enrollment and employment. To identify the impact of having more upper-income (or higher-achieving) peers, I use the within-school residual variation from school trend in the proportions of upper-income students. The residual cross-cohort variation in the proportion of upper-income peers is unlikely to impact students' access to school resources (as I show empirically) and so can isolate the effect of peer income independent of access to school

---

<sup>1</sup>Social capital theory has been applied in education and employment to identify the drivers of inequities (e.g., Coleman, 1988; Fernandez and Fernandez-Mateo, 2006; De Giorgi, Pellizzari, and Redaelli, 2010; Schmutte, 2015).

resources. I also disentangle the role of peer income from peer achievement by controlling for the share of higher achieving peers.

I draw two main conclusions: first, lower-income students are exposed to very few peers with higher family income (upper-income students). The gap in exposure to upper-income classmates between lower- and upper-income students is approximately double the gap in exposure to higher-achieving classmates—41 percentage points relative to 20 percentage points. Classroom assignment contributes very little to the low exposure to upper-income peers. The low exposure to upper-income peers is mainly driven by residential and school segregation. Second, the positive peer income effect on lower-income students’ college enrollment and employment does not seem to operate through peer achievement. Prior work on peer effects in k-12 has generally focused on peer academic achievement (see, e.g., Imberman et al, 2012; Lavy, Paserman and Schlosser, 2012; Jackson, 2013; Feld and Zoelitz, 2017).

To capture the likelihood that lower-income students are exposed to upper-income peers in their classrooms, I use Texas state administrative data from 2004 to 2022 in which I have information on the exact classrooms students are in and follow students from grades 5 to 12. I define exposure as the proportion of a student’s cumulative number of classmates between grades 5 and 12 who are higher income (excluding the student herself). I use “higher/upper income” and “lower income” as shorthand for students who are always below the free/reduced lunch income eligibility cutoff (income below \$51,338 for a family of four) and those who are always above the free/reduced lunch eligibility cutoff, respectively. Upper-income students constitute approximately 24% of students enrolled in public schools in Texas.<sup>2</sup> For comparability, I define higher-achieving students as those who scored in the top 24 percentiles of reading test scores in grade 4.<sup>3</sup>

On average, 10% of lower-income students’ classmates are upper income over the period from grades 5 to 12 compared to 51% of upper-income students’ classmates being upper-income. The distribution of exposure is right-skewed: the median lower-income students’ proportion upper-income classmates is 6% and one in every five lower-income students goes through grades 5 to 12 with a share of upper-income classmates smaller than 1%.<sup>4</sup> Lower-income students are more likely to be exposed to higher-achieving students (students who scored in the top 24 percentiles of their grade 4 reading tests) than they are to be exposed to upper-income

---

<sup>2</sup>The income cutoffs described are based on free/reduced lunch income eligibility in 2022. This definition of economic disadvantage builds on Michelmore and Dynarski (2017), who find that the number of years on free/reduced lunch captures student economic disadvantage better than a binary measure of economic disadvantage based on one year of free/reduced lunch eligibility status. Among students for whom financial aid data are available, I find that those with lower and higher incomes have a median parental income of approximately \$22,000 and \$117,119, respectively.

<sup>3</sup>Test scores are based on standardized grade 4 TAKS (2007–2011) and STAAR (2012–2018) reading tests. I use grade 4 because it is prior to when I start following students in the data and documenting their cross-income and test-score exposure. The estimates are similar if I instead used their math test or their average standardized score between grades 3 and 8.

<sup>4</sup>Students on average have 2000 classmates between grades 5 and 12; i.e., 20% of lower-income students go through grades 5 to 12 interacting with fewer than 20 higher-income classmates. I equally weight every time a student shares a class with another student.

students. On average, 18% of lower-income students' classmates are high-achieving over the period from grades 5 to 12. The gap in exposure to higher-achieving students is consistent if instead achievement is based on average grades 3 to 8 reading scores or math test scores.

The low exposure to upper-income students is mainly driven by residential and school socioeconomic segregation. Residential (district) segregation and school (within district) segregation each account for 71% and 23% of the difference between lower-income students' exposure to upper-income peers and the socioeconomic integration benchmark, respectively. For the average lower-income student, within-school classroom sorting by income does not matter much for exposure to upper-income students: had students been randomly assigned to classrooms, their exposure to upper-income students would have increased by only 0.7 percentage points. The relatively small contribution of classroom assignment suggests policies seeking to improve cross-income exposure should focus on residential and school segregation. However, the average contributing role of the classroom to exposure to upper-income students misses important student heterogeneity in the potential gains from classroom integration: at the 10th percentile, lower-income students would have been exposed to 4.2 percentage points more upper-income students had students been randomly assigned to classrooms. The relative importance of residential segregation is consistent with the findings of Owens, Reardon, and Jencks (2016), who report that approximately two-thirds of income segregation between schools is due to segregation between districts. However, Owens, Reardon, and Jencks (2016) did not have classroom level data and so were not able to quantify the role of the classroom.

Lower-income students' school exposure to upper-income classmates is highly correlated with the likelihood of befriending upper-income students. To identify the relationship between exposure to upper-income peers and cross-income friendship measures related to economic mobility, I link publicly available school-level data published by Chetty et al. (2022) on the likelihood of cross-income friendships with school measures of cross-income exposure in Texas. I find that low-SES students' school average proportion of high-SES friends (school economic connectedness as captured by Chetty et. al (2022)) is highly correlated with lower-income students' school average proportion of upper-income classmates in Texas as shown in Figure 11: the correlation is approximately 0.9.<sup>5</sup> I also find that classroom assignment by income within schools in Texas is correlated with lower-income students' likelihood of befriending high-SES students conditional on school composition (school friending bias) as measured by Chetty et al. (2022): the correlation is approximately 0.6. The high correlation between within school sorting and friending bias suggests that school differences in cross-income friendships may in part be explained by differences in students' classroom assignment by income. However, since classroom assignment contributes very little to overall exposure to upper-income

---

<sup>5</sup>The economic connectedness measure is based on dividing the average proportion of high-SES friends by the the proportion of high-SES students in the population.

peers it might be better to focus on district and school integration policies.

Having established that there are high rates of income segregation and that exposure to upper-income students relates to cross-income friendship formation, the question that follows is whether exposure to upper-income students matters for lower-income students' college enrollment and wages. To address selection concerns and try to isolate the role of peers independent of resources, I use cross-cohort within-school variation in exposure to upper-income students. I find that a 10-percentage-point increase in the proportion of upper-income students increases enrollment in 4-year public colleges by 0.38 percentage points, quarterly wages at 22–25 by 1.7%, and income percentile rank by 0.33.

The impact of the share of upper-income peers on college enrollment and wages is not sensitive to controlling for the share of higher achieving peers, suggesting that the impact of peer income on lower-income students' college enrollment and wages is driven by things orthogonal to students' peer academic achievement. It seems that having more higher-achieving peers has a negative impact on lower-income students' likelihood of enrolling in college and wages. Feld and Zöllitz (2017) using random assignment to sections at university similarly find that peer academic achievement may negatively impact lower-achieving students.<sup>6</sup>

The relationship between exposure to upper-income peers and long-term outcomes may be through lower-income students accessing additional school resources when sharing a classroom with upper-income students. Estimates of the impact of having a higher proportion of upper income peers on college enrollment and wages are based on within-school temporary deviations from school trend in the proportion of upper-income peers. These temporary (small) changes in a cohort's proportion of upper-income peers are less likely to impact students' access to school resources (e.g., teacher quality, AP courses offered, and school budgets). Consistent with this, I find that residual variation in peer income has no impact on lower-income students' proportion of novice teachers, number of AP courses taken, and average school per-pupil spending, suggesting that the estimates are not driven by changes in access to school resources. However, the relationship between exposure to upper-income students and long-term outcomes may be driven by other changes in access to school resources that are not observed in administrative data (e.g., higher-income parents volunteering for school events).

This paper contributes to several literatures. First, it contributes to the segregation literature. This literature has focused primarily on racial/ethnic segregation (e.g., Clotfelter, Ladd, and Vigdor, 2002; Clotfelter, Ladd, Clifton and Turaeva, 2021) and academic sorting within schools (e.g., Antonovics, Black, Cullen, and

---

<sup>6</sup>In Texas, the 10% rule—students who performed in the top 10% of their high school are automatically admitted to state universities—may influence the impact of an increase in the proportion of higher achieving peers on students' long-term outcomes. That said, Black, Denning, and Rothstein (2022) find no evidence of a negative impact on college enrollment and earnings of high achieving students in more advantaged schools who did not benefit from the 10% rule.

Meiselman, 2022; Lucas and Berends, 2002). I study the implications of segregation for cross-income exposure and how it relates to long-term outcomes. This prior literature is subject to multiple data limitations that hamper researchers' ability to document the effects of segregation on exposure and long-term outcomes, including a lack of observations on students within a classroom and/or on their college enrollment and employment outcomes (Owens, Reardon and Jencks, 2016; Dalane and Marcottee, 2020; Kalogrides and Loeb, 2013). Using rich administrative data from Texas, I can follow k-12 students over time through early adulthood, observe their classroom enrollment and peers, and link their classroom peer composition to long-term outcomes. Second, my work contributes to the literature on social capital and economic mobility. I build on the work of Chetty et al. (2022b) on the relationship between cross-income friendships and long-term outcomes by identifying the source of the variation in exposure to upper-income peers and disentangling the impact of peer income from the effects of peer test scores and access to resources.

Third, this paper relates to the large peer effects literature examining the relationship between peer composition and short- and long-term outcomes (examples include Sacerdote, 2001, 2011; Zimmerman, 2003; Lavy, Paserman and Schlosser, 2011; Black, Devereux and Salvanes, 2013; Cools, Fernandez and Patacchini, 2019; Zimmerman, 2019; and Rao, 2019). It is most closely related to the paper by Cattan, Salvanes and Tominey (2022), which investigates the impact of exposure to the children of parents who attended elite schools in Norway. They find that exposure to elite peers increases enrollment in selective universities. I focus on the impact of exposure to upper-income students (regardless of parental education) on teacher quality, advanced coursework, college enrollment, and wages. Importantly, I examine whether the impact on college enrollment and wages is driven by having more upper-income peers, independent of their academic achievement and/or access to school resources.

The paper is organized as follows. In Section 2, I begin by laying out the theoretical framework of the analysis. Section 3 defines my measures of cross-income exposure and data used. Section 4 reports estimates of cross-income exposure at the state level and how they compare to cross-test score exposure, as well as district, school and classroom integration benchmarks. Section 5 describes the relationship between exposure to upper-income students and friendship formation. Section 6 reports the relationship between cross-income exposure and long-term student outcomes. Section 7 concludes the paper.

## 2 Conceptual Framework

Exposure to higher-income peers is a function of the district a student resides in, the school they enroll in, and the classroom they are assigned to or choose. At each of these levels, parents and students make decisions

about where to live and what school and classroom to enroll in given their economic resources (household income) and subject to potential information imperfections and discriminatory barriers. These decisions and constraints shape not only their exposure to higher-income peers but also their access to neighborhood and school resources. The first part of the paper attempts to draw a complete picture of how each of these layers (district, school, and classroom) contributes to lower-income students' exposure to higher-income peers across grades.

Theoretically, exposure to higher-income peers could shape lower-income students' long-term outcomes in at least two ways: resource allocation and/or social capital.<sup>7</sup> The distribution of students by income across and within schools could shape how financial and instructional resources are allocated (e.g., Owens and Candipan, 2019 and Kalogrides and Loeb, 2013). Owens and Candipan (2019) find that schools in higher-income neighborhoods tend to have higher spending per student and higher salaries for teachers. Unequal resource allocation can also take place within a school, between classrooms. Kalogrides and Loeb (2013) find that higher-income students tend to be assigned more experienced teachers than lower-income students in the same school.

Social capital has long been discussed as a type of capital that is independent of human and physical capital that may impact productivity (Coleman, 1988). Social capital is defined by the structure of relationships between individuals (Coleman, 1988). There are two potential channels through which social capital may impact students' long-term outcomes: information and norm setting (Coleman, 1988). An example of an information channel is if higher-income students are more likely to know which courses to take to improve their likelihood of being accepted into more selective colleges and that they are more likely to share this information with lower-income students when sharing the same classroom. Bourdieu (1997) suggests that individuals from higher-income households have higher cultural capital, which results in the persistence and reproduction of social structures. Cultural capital is defined as knowledge, skills, and tastes that are acquired because an individual belongs to a social class. The information channel is likely weaker the more information and skills are available to students through formal sources, for example through counselors.

The social norms channel depends on the prevalence (actual/assumed) of a behavior in a group. For example, if higher-income students are more likely to enroll in selective universities, this may create a norm/expectation that students apply and enroll in selective universities across income groups. Several papers have documented the potential importance of social norms on academic performance in schools, perceptions, and giving behavior (Frey and Meier 2004; Bursztyn and Jensen, 2015; Bursztyn, González,

---

<sup>7</sup>Other potential channels include classroom spillover and class rank. For example, changes in peer income composition may impact a teacher's ability to target instruction by changing how homogeneous a classroom is. Changes in classroom peers may also impact a student's rank in the class, which may have implications for her aspirations and long-term outcomes (Cattan, Salvanes, and Tominey, 2022).

and Yanagizawa-Drott, 2020).

The strength of the information and norm channels depends on the school and social structure. If high-and low-income students enrolled in the same school are likely to attend different classrooms (because of high sorting by income/test score), then they are less likely to pass on information and shape cross-group norms. It also depends on the strength of the cross-group relationships. If students are less likely to form friendships across income categories, then they may be less affected by cross-income exposure. Hoxby (2000) finds that peer achievement effects are stronger within a race. Michelman, Price and Zimmerman (2022) find that exposure to “high-status” students at Harvard in the 1920s and 1930s disproportionately improved access to elite social clubs for students who graduated from private schools. Similarly, Cattan, Salvanes, and Tominey (2022) find that the impact of having more classmates whose parents graduated from elite schools is stronger among high-SES students than among low-SES students. Section 5 examines whether and how changes in cohort composition impact the level of classroom sorting by income and interacts with student characteristics (student test scores and race/ethnicity).

The social capital channel may operate exclusively through students’ academic achievement. In other words, it could be that upper-income students happen to perform better academically, and being around higher achieving students—*independent* of their income—has an impact on lower-income students’ academic performance and labor outcomes. Some of the channels through which a higher proportion of higher achieving peers may impact student outcomes include: student class rank, teacher behavior and/or academic spillover through peer-to-peer tutoring. The peer effects literature suggests the impact of peer achievement may vary depending on students’ own baseline academic performance (Cools, Fernández and Patacchini, 2019; Feld and Zöllitz, 2017, Lavy, Paserman and Schlosser, 2011; Sacerdot, 2011).

### **3 Data and Measures of Cross-Income Exposure**

#### **3.1 Data**

I use longitudinal administrative data from the Texas Education Research Center (ERC) that link student data from the Texas Education Agency (TEA) to Texas Higher Education Coordination Board (THECB) and Texas Workforce Commission (TWC) data. These TEA data cover the period from 2004 to 2022 and include student test scores, courses (and class assignment starting in 2011), demographics, attendance, graduation, and assigned teachers (including teacher certification and demographics). My main measure of income is free/reduced lunch status. I use years of assignment to free/reduced lunch status to capture the degree of

economic disadvantage.<sup>8</sup>

The college enrollment data (THECB) are from 2008 to 2022, but I focus on the 2014 to 2022 data. The THECB data comprise information on college applications to state public schools (including information on parental income and education status), as well as financial aid data, which also includes information on parental income. I focus on the following college enrollment outcomes: any college enrollment, 2-year public college enrollment, 4-year public college enrollment, highly selective college enrollment, and graduation from any college. Colleges are defined as selective if they are between levels 1 and 4 based on Barron's selectivity index (most to very competitive in 2009). The college data are limited to enrollments in the state of Texas.

The employment data (TWC) are from the years 2008 to 2022. These data include information on industry, employment level, and wage. Quarterly wage data are based on unemployment insurance data and hence cover only employment in Texas. In Section 5, I use quarterly wage data from 2022. For each student, I capture the average quarterly wage from employment in the first three quarters of 2022. I use both the average quarterly wage and the percentile-ranked wage (ranked within cohort).

To identify and better understand the link between exposure and friendship formation, in Section 5, I supplement the analysis with two other sources of data: Add Health data and publicly available high school measures of the likelihood of cross-income friendship based on Facebook data from Chetty et al. (2022), both of which I discuss in more detail in Section 5.

### 3.2 Measure of Family Income

To measure student income, ideally, I would have parental income for every enrolled student. In the absence of data on parental income, I use the proportion of years on free/reduced lunch to capture the degree of economic disadvantage.<sup>9</sup> I divide students into three main income groups: always, sometimes and never on free/reduced lunch.

The segmentation of students based on proportion of years on free/reduced lunch seems to capture the variation in parental income as shown in Figure 1. For students with financial aid data and who are never, sometimes and always in free/reduced lunch status, their average adjusted parental income is \$141,686, \$51,406, and \$27,305, respectively.<sup>10</sup> I define higher-income students as those never in free/reduced lunch

---

<sup>8</sup>Test scores are mainly based on standardized grade 4 and 8 TAKS (2007–2011) and STAAR (2012–2018) reading and math tests.

<sup>9</sup>To identify the proportion of years on free/reduced lunch, I use all years from 2004 to 2022.

<sup>10</sup>Median parental income is \$117,119, \$39,145 and \$22,000, respectively. Students with financial aid data are a select, systematically different, group. Therefore, these averages are not representative of the average income for each of the groups. It provides us with a general sense of the variation and is likely an upper bound for students who are always on free/reduced lunch. Financial aid data are available for 52%, 35% and 30% of students who are never, sometimes, and always in free/reduced lunch status, respectively.

status. Higher-income students represent approximately 24% of the student population. Students sometimes and always in free/reduced lunch status constitute approximately 48% and 29% of the student population, respectively.

Using the proportion of years on free/reduced lunch seems to effectively capture the degree of economic disadvantage. Among students with financial aid data, I find that 59% of students who are always in free/reduced lunch status reported parental income that is below the federal poverty cutoff for a family of four.<sup>11</sup>

I test whether these patterns hold using other measures of income. For students who apply to public universities in Texas, I have additional information on parents' income bracket.<sup>12</sup> Of students who applied to public universities, 81% 24% and 3% who were never, sometimes and always on free/reduced lunch, respectively, report having parental income above \$80,000.

For the rest of the paper, I refer to students who are never on free/reduced lunch as higher- or upper-income. I refer to students who are sometimes and always on free/reduced lunch as middle- and lower-income students, respectively. This paper focuses on higher- and lower-income students.

### 3.3 Measure of Cross-Income Exposure

I define cross-income exposure as the proportion of classmates of another income group in a student's classroom. I focus on the proportion of higher-income students. I capture the cumulative exposure to higher-income students by following three cohorts of students (2012–2014) starting from fifth to expected twelfth grade in Texas. In each year, I capture the number of classmates in a student's classroom (excluding self). I also capture the number of higher-income classmates in the classroom (excluding own status). I divide the total number of higher-income peers student  $i$  encountered in all the classrooms they were enrolled in between grades five and twelve ( $N_{(-i)}^H$ ) by the total number of peers student  $i$  encountered in those classrooms ( $N_{(-i)}$ ): 
$$\frac{N_{(-i)}^H}{N_{(-i)}}.$$

One can understand this measure as capturing the proportion of potential interactions rather than the proportion of students. If a student is in multiple classrooms with the same student, the latter is included in the denominator (total student interactions) multiple times. She would also be included in the numerator if she were of higher income multiple times. The assumption is that what matters is the proportion of interactions rather than the proportion of students. Each interaction is likely to yield an additional benefit.

---

<sup>11</sup>Based on reported federal poverty cutoffs for 2020 to 2022.

<sup>12</sup>I have this information for 33% of students who are never on free/reduced lunch and only 18% and 14% of students who are sometimes and always on free/reduced lunch.

I assume that each interaction has an equal level of benefit.<sup>13</sup>

## 4 How Likely is Cross-Income Exposure and What is the Contribution of District, School and Classroom Economic Segregation?

### 4.1 Sample

To answer this question, I follow three cohorts of students from grade 5 to expected grade 12. The three cohorts are the 2012, 2013 and 2014 cohorts. Students enrolled in any public school (including charter schools) in Texas in grade 5 in academic year 2012 belong to the 2012 cohort. These cohorts of students are followed in the data, and their classroom income composition is documented every year from 2012 to 2019. Across these cohorts, there are 1,128,554 students. Table 1 summarizes the characteristics of the sample. Approximately 50% of the cohort identify as Latin-American students, 14% as Black students, and 68% as White students. Most of the students are enrolled in traditional public schools—4–6%, in a given grade, are enrolled in a charter school, as shown in Table 10.5. Table 10.5 summarizes the enrollment patterns of the students in each expected grade. Note that “expected” grade is not their actual grade but the grade in which we expect them to be enrolled given the year and their cohort.

### 4.2 Main Patterns of Exposure to Upper-Income Students

I find that lower-income students between grades 5 and 12 are one-fifth as likely to be exposed to higher-income students than their higher-income counterparts.<sup>14</sup> Figure 2 shows that the proportion of higher-income classmates in the classrooms of higher, middle and lower income students is 51%, 19% and 10%, respectively.<sup>15</sup>

The distribution of exposure to higher-income peers is skewed to the right for low-income students, as

<sup>13</sup>Students may be in different classrooms with the same student. The second time they are with the same student, the gain from exposure to that student may be nonlinear. Similarly, the time spent with students in the same classroom can influence the amount of gain they would reap from that classroom exposure. Future work should identify whether the gain from classroom exposure to the same student and/or the length of sharing a classroom influences the expected gain from that interaction. I assume equal gain from each classroom exposure independent of the length of sharing a classroom or the number of classrooms shared with the same student. Note that if higher-income students systematically enroll in more classrooms, the average proportion of interactions with higher-income students may be higher than the average proportion of higher-income students.

<sup>14</sup>The difference in cross-income exposure can also be presented as differential exposure to lower-income students. The proportion of lower-income classmates for higher- and lower-income students is 12% and 44%, respectively. In Appendix Figure A3, I present the average exposure to students with lower incomes by income group.

<sup>15</sup>If we instead segment students by years on free/reduced price lunch, we see that exposure to higher-income peers decreases as the number of years in free/reduced price lunch status increases as shown in Figure A2. The gap in exposure to higher-income students is similarly large if we define exposure to higher-income students as exposure to students who have financial aid data and whose parental gross income is greater than \$150,000 or \$200,000 based on financial aid data, as shown in Figure A1.

shown in Figure 6 Panel (a). A large proportion of students go through grades five to twelve exposed to a small number of higher-income students. One in every five lower-income students goes through grades 5 to 12 interacting with fewer than 20 higher-income students, and less than 1% of their classmates are higher income.<sup>16</sup>

The difference in exposure to upper-income students is a function of residential, school, and classroom socioeconomic segregation. To identify how each of residential economic segregation, school segregation and classroom assignment contribute to exposure to upper-income peers I compare students' exposure to upper-income students to three benchmarks. In the first benchmark, students enrolled in the state are randomly assigned to districts, schools, and classrooms (district integration benchmark). In the second, the district composition is fixed, and students are randomly assigned to schools and classrooms (school integration benchmark). In the third, the school composition is fixed, and students are randomly assigned to classrooms (classroom integration benchmark). In practice, this is very similar to comparing lower-income students' average exposure to upper-income students to the proportion of upper-income students in the state, district and school. For example, if we were to randomly assign students across districts, the expected exposure to higher-income students would be equal to the proportion of higher-income students enrolled in the state of Texas.<sup>17</sup>

I show these benchmarks in Figure 2 with students' actual exposure to upper-income classmates. The difference between actual exposure and the district benchmark for full integration is 11.6 percentage points. In other words, had students been randomly assigned to districts, lower-income students would have been in classrooms with 11.6 percentage points more upper income classmates—that is more than twice lower-income students' current exposure. The difference between actual exposure and the school benchmark for full integration is 3.4 percentage points, and the difference between actual exposure and the classroom benchmark for full integration is 0.7 percentage points. The relatively large difference between the district and school integration benchmarks is consistent with the findings of Owens, Reardon, and Jencks (2016), who report that approximately two-thirds of income segregation between schools is due to segregation between districts. However, Owens, Reardon, and Jencks (2016) did not have classroom level data and so were not able to quantify the role of the classroom.

---

<sup>16</sup>On average, students during this period encounter 2000 classmates. Some may be repeated classmates in different classrooms.

<sup>17</sup>Since I am excluding own status, the expected exposure for higher- and lower-income students under random assignment will be slightly different; higher-income students would be exposed to fewer higher-income peers since their own status is not included in the numerator. This matters more in small units such as school classrooms. In each year, I calculate the benchmarks based on the number of students in the defined cohort still enrolled in a public school that year, as well as any students that share a classroom with a student in the cohort. This is particularly important in high school, where students may be enrolled in classrooms with students in other cohorts, and so their potential exposure may be different. I present my calculation of expected exposure under random assignment—excluding students' own status—in Appendix (A).

The small difference between actual exposure and the classroom integration benchmark suggests that, for the average lower-income student, classroom assignment accounts for a very small portion of the low exposure to upper-income students. However, school classroom-level policies may be easier to implement, and popular district-level policies such as the addition of charter schools appear to have an even smaller impact on cross-group exposure.<sup>18</sup> The average contributing role of the classroom to exposure to upper-income students also misses important student heterogeneity in the potential gains from classroom integration: at the 10th percentile, lower-income students would have been exposed to 4.2 percentage points more upper-income students had students been randomly assigned to classrooms. I find that 67% of lower-income students, compared to 30% of upper-income students, are exposed to fewer upper-income students than would be expected had students been randomly assigned to classrooms. In other words, a school's proportion of upper-income students overestimates lower-income students' exposure to upper-income students, on average.

The difference between actual exposure and the classroom integration benchmark can be explained by students taking different courses and/or students taking the same course but enrolling in different classrooms. In Figure A11, I add an additional benchmark: expected exposure in school course to upper-income students under random assignment to classrooms and label it “random class in course”. The difference between actual exposure and the “random class in course” benchmark is 0.5 percentage points; most of the difference between actual exposure and the classroom exposure benchmark can be explained by students taking the same course, enrolling in different classrooms.

The large contribution of residential and school socioeconomic segregation on students' exposure to upper-income peers would prompt us to think of policies that focused on greater district and school integration (e.g., racial integration policies such as Metropolitan Council for Educational Opportunity (METCO)'s voluntary busing program in Boston and race-based busing in Charlotte-Mecklenburg schools).<sup>19</sup> However, district (school) integration policies may prompt districts (schools) to track students by income into different schools (classrooms).

To capture how the role of school and classroom sorting may vary by district composition, I split districts into percentiles based on the proportion of higher-income students in the district. Then, I present the average exposure of lower-income students to upper-income students in each district percentile and how it compares to the school and classroom full-integration benchmarks. Figure 3 captures the percentage point difference

---

<sup>18</sup>Monarrez, Kisida, and Chingos (2022) find that a charter school opening decreases the exposure of Black and Hispanic students to White and Asian students by approximately 0.3 and 0.1 percentage points, respectively.

<sup>19</sup>Setren (2024) finds that METCO improved bussed minority students' test scores and college enrollment and that bussed students were tracked in receiving schools to lower-performing classes. Bergman (2018) finds that a lottery-based (voluntary) desegregation program had a mixed impact on participating minority students who received offers to attend schools serving high-income, predominately white students; participating students were both more likely to enroll in college and be arrested for nonviolent offenses. Billings, Deming, and Rockoff (2014) find that the end of race-based busing widened racial gaps.

between the exposure of students to upper-income classmates and the school and classroom integration benchmarks. The school integration benchmark is captured by the height of the navy bar, and the classroom integration benchmark is captured by the height of the gray bar. The exposure of students to upper-income students in each percentile is captured by the height of the light blue bar.

I find that schools (and districts) that are more integrated do not fully capture the benefit of having a higher proportion of upper-income students. The difference in exposure to upper-income students between the school and classroom benchmarks for integration and actual exposure are 13 and 2.6 percentage points in higher-income districts (districts enrolling 30% or more upper-income students), respectively.<sup>20</sup> The importance of classroom integration, particularly in higher-income districts, becomes increasingly important in high school as shown in Figure A12.<sup>21</sup>

If race and ethnicity are highly correlated with income, in practice, what appears to us to be socioeconomic segregation may actually be capturing racial/ethnic segregation. In fact, as shown in Figure A4, higher-income students account for only 9% and 11% of Hispanic and Black students, respectively, while they constitute the majority of White students (51%). To disentangle cross-racial/ethnic from cross-income exposure, I first present exposure to upper-income students by student race/ethnicity. Then, I capture exposure to same-race/ethnicity upper-income classmates.

Lower-income students in all racial/ethnic groups are exposed to fewer upper-income students than their higher-income counterparts as shown in Figure 4.<sup>22</sup> The difference in exposure to upper-income classmates is smaller within racial/ethnic groups, suggesting that part, but not all, of the difference in exposure is a function of racial/ethnic segregation as shown in Figure A5. For Hispanic students, 28% and 4% of higher- and lower-income students' Hispanic classmates are higher-income. For Black students, 28% and 7% of higher- and lower-income students' Black classmates are higher income. For White students, 66% and 32% of higher- and lower-income students' White classmates are higher income.<sup>23</sup>

<sup>20</sup>The increasing gap between actual exposure and the school and classroom benchmarks is partly mechanical. There is more potential for sorting as the proportion of higher-income students increases. Although part of this may be mechanical, schools/districts with a higher proportion of higher-income students also have the highest potential to expose lower-income students to more upper-income peers, and the gap in exposure captures that missed potential. Of lower-income students, 12% reside in districts containing 30 percent or more upper-income students.

<sup>21</sup>Sorting across courses by income becomes increasingly important in high school, as shown in Figure A12. In grade 9 half of the difference between actual exposure and the classroom integration benchmark is driven by upper- and lower-income students taking different courses. The increasing importance of classrooms in older grades is consistent with Clotfelter, Ladd, Clifton and Turaeva (2021), who find that within-school segregation by race/ethnicity accounts for a larger portion of total segregation in high school.

<sup>22</sup>White students across incomes are exposed to approximately 10 percentage points more upper-income classmates.

<sup>23</sup>The low exposure of Hispanic and Black, low-income students to upper-income classmates of the same race/ethnicity is concerning if peer effects are stronger within racial groups. Students who share multiple characteristics may be more likely to befriend one another (Moody, 2001; Tuma and Hallinan, 1979). Hoxby (2000) finds that peer achievement has a stronger impact on students' test scores within races.

### 4.3 Exposure to Upper-Income Students Compared to Exposure to Higher-Achieving Students

The difference in exposure to upper-income students may capture differences in exposure to higher-performing students. Understanding how the relationship between exposure to upper-income students compares to exposure to higher-achieving students has practical and policy implications. If exposure to upper-income students is analogous to exposure to higher-achieving students, then it would be difficult to disentangle the spillover from having upper-income peers from having higher-achieving peers, and income desegregation policies may address both gaps in exposure to upper-income and higher-achieving students. If instead peer income and achievement compositions capture two different phenomena, then we could identify the impact of each on students' long-term outcomes separately, and interventions that address gaps in peer achievement would be different from those that address gaps in peer income.

I identify how exposure to upper-income students compares to exposure to higher-achieving students. Higher-achieving students are defined as students who scored in the top 24 percentiles of reading test scores in grade 4. Then, I follow students from grade 5 to grade 12 and document their exposure to higher-achieving students (instead of upper-income students). I chose the highest 24% of academic performance so that it is comparable to the proportion of students who are upper income (24% of students are upper income). I find similar patterns if I define higher-achieving students by their grade 4 math test scores.<sup>24</sup>

The difference in exposure to higher-achieving students between upper- and lower-income students is considerably smaller than the difference in exposure to upper-income students. As shown in Figure 5, lower-income students are exposed to 20 percentage points fewer higher-achieving students than higher-income students. In comparison, the gap in exposure to upper-income students is closer to 41 percentage points. The gap in cross-test score exposure is also smaller than cross-income exposure as shown in Figure A8.<sup>25</sup> The smaller difference in exposure between high- and low-achieving students suggests that students are less clustered by test scores than by income.

There are more lower-income students exposed to very few upper-income students than those exposed to very few higher-achieving students. The distribution of exposure to upper-income students is right-skewed, but exposure to higher-achieving students is normally distributed as shown in Figure 6 Panel (b): Lower-income students at the 25th percentile of the peer income distribution are exposed to 1.4% upper-income

---

<sup>24</sup>I use grade 4 because it is prior to when I start following students in the data and documenting their cross-income and test-score exposure and more recent scores may be more predictive of long-term outcomes. That said, one test score may be subject to larger measurement error compared to the income measure which is based on multiple years. To address the concern with measurement error I present the results using grades 3 to 8 average test scores as well.

<sup>25</sup>Lower-achieving students are those who scored in the bottom 29 percentiles of their grade 4 reading test.

classmates. In comparison, lower-income students at the 25th percentile of the peer achievement distribution are exposed to 12% high-achieving classmates.

I use grade 4 because it is prior to when I start following students in the data and more recent scores may be more predictive of long-term outcomes. That said, one test score may be subject to larger measurement error compared to the income measure which is based on multiple years of FRPL status. To test whether the patterns are driven by how noisy the measure of achievement is, I instead define high-achieving students as those who score in the top 24 percentiles of the average grades 3 to 8 reading scores. The gap in exposure to high-achieving students between upper- and lower-income students is similar at 23 percentage points if achievement is defined by average grades 3 to 8 reading tests as shown in Figure A9.

#### 4.4 Variation in District and School Exposure to Upper-Income Students, Conditional on Student Income Composition

Low exposure to upper-income students is a function of both who enrolls in a district/school and how districts and schools are organized. Two districts (schools) could have the same proportion of upper-income students, but in one district (school) lower-income students are exposed to fewer upper-income students because they attend different schools (classrooms).

I use the variance ratio to capture the exposure gap conditional on the proportion of higher-income students served in a school or district. Unlike the exposure measure, the variance ratio accounts for the change in the range of the potential gap between actual and expected exposure under random assignment by dividing the numerator by the range of potential exposure as shown in Equation (6) in Appendix B, where the variance ratio is discussed in more detail. In Figures 7 and 8, I plot the exposure of lower-income students to higher-income students in the district (school) against the proportion of higher-income students in the district (school).

In districts in particular, we can see that there is substantial variation in lower-income students' exposure to higher-income students among districts with similar compositions of students, which is explained by differences between school sorting levels. In districts with high levels of between-school sorting (top 20th percentile), lower-income students are exposed to 6.8 percentage points fewer upper-income classmates than the district average.<sup>26</sup> A factor that could explain differences in exposure to upper-income students,

---

<sup>26</sup>I map out districts' between- and within-school sorting rates in Figure A17. I also show the average difference between observed and school and classroom integration benchmarks in Figure A16. There are several districts in the west of Texas that have relatively large gaps in exposure to higher-income students and a lower proportion of higher-income students. There are also areas in the south of Texas where a high proportion of higher-income students reside, and the gap in cross-income exposure is smaller.

conditional on district income composition, is the number and type of schools offered. I find that districts with more charter and private schools relative to the number of students served tend to have higher levels of between-school sorting by income, as shown in Figure 9. Refer to Appendix C for more details on the method and findings.

In schools, school composition is highly predictive of the exposure of lower-income students to higher-income peers, but there is still some difference in the likelihood of being exposed to higher-income students, captured by the variation in the level of within-school sorting. In schools with high levels of within-school sorting (top 20th percentile), lower-income students are exposed to 2.5 percentage points fewer upper-income classmates than the school average. I find that the within-school sorting by income (variance ratio) declines by approximately 1.5 percentage points, from 11.5 to 10, when I control for students' grade 8 test scores and another 0.6 percentage points when I add an indicator for whether the student has taken Algebra I by grade 8. This suggests that academic preparation explains part, but not all, of the difference in classroom sorting by income within schools.<sup>27</sup>

The number and type of courses offered in a school could impact how students are tracked by test score and income. I find that the number of science and math courses offered in a school correlates most strongly with classroom sorting by income, regardless of whether those courses are advanced. This is shown in Figure 10.<sup>28</sup> These patterns may reflect schools' tracking policies or other school and student characteristics that correlate with sorting by income. In other work, I examine the impact of the addition of an advanced course using variation in the timing of when an advanced course is added to a subject area. I find that the addition of an AP course increases exposure to upper-income students and it appears to be driven by an increase in the proportion of students who take an AP course in the subject area (Mallah, 2024).

## 5 How Does Cross-Income Exposure Relate to Friendship Formation?

In prior work Chetty et al. (2022) find that cross-income friendships are strongly related to prospects of economic mobility. Differences in patterns of cross-income friendships may be attributed to differences in exposure to upper-income peers in schools and classrooms. In this section, I examine whether and how

<sup>27</sup>In Table 2 Model (4), I find that the variance ratio declines only slightly (by 0.08 percentage points) when I also control for student attendance and suspension records in expected grade 8, which suggests that these "behavioral" controls do not provide any additional information. This estimate is based on regressing an indicator for student income on a student's proportion of upper-income classmates in a school. I include school and cohort-year fixed effects so that the coefficient on student income captures the difference between the proportion of upper-income students in upper-income relative to other students' classrooms in the same school.

<sup>28</sup>I regress the level of within-school sorting by income (captured using the variance ratio) for high schools on the number of courses offered divided by the number of students enrolled in 2019. The variance ratio is standardized based on the distribution of within-school sorting in schools weighted by the number of students enrolled. Similarly, the ratio of courses to students is standardized based on the distribution of courses to students in the sample of high schools in 2019, weighted by the number of students enrolled.

school exposure to upper-income peers relates to the likelihood of forming cross-income friendships.

A limitation of the Texas data is that they lack information on cross-income friendship formation. As such, I supplement the analysis with two other sources of data: Add Health data and publicly available high school measures of the likelihood of cross-income friendship based on Facebook data from Chetty et al. (2022). The Add Health data have two advantages over the Texas data. First, they come from a nationally representative sample of students in grades 7–12; second, they contain information on individual student friendship patterns and enrollment in extracurricular activities. However, they lack information on which classrooms students are enrolled in, and as such, I use it to capture the relationship between school (and extracurricular activities) peer income composition and friendship formation. The in-school survey asks students to list up to five male and five female friends. Table A1 summarizes the demographics of students in Add Health.<sup>29</sup>

To capture the relationship between school (and extracurricular activities) peer income composition and friendship formation, I calculate the proportion of each student's listed friends who are high SES in the Add Health data. I define high-SES students as those in the top 24 percentiles of predicted income. I use student's predicted household income instead of actual household income because I only have information on household income for a subset of students surveyed.<sup>30</sup> Then, I calculate the proportion of a school's student population that is high-SES and the proportion of a student's peers in extracurricular activities that are high-SES (excluding herself). Last, I run a regression with the proportion of high-SES peers in school (and another with the proportion of high-SES peers in extracurriculars) on the student's proportion of high-SES friends.

I find that exposure to high-SES peers in school and extracurricular activities is highly predictive of a student's friendship patterns. I find that for low-SES students, both the school and extracurricular income compositions are strongly correlated with the proportion of high-SES friends: correlations of 0.37 and 0.40, respectively.<sup>31</sup> The relationship between exposure to high-SES students and friendship formation is presented in Figure 12. The fitted line is below the 45 degree line, suggesting that the composition of the school and extracurriculars does not perfectly predict the average difference in the proportion of high-SES peers (it is not 1-to-1) for low-SES students, but it comes close: For low-SES students, the slope between school

---

<sup>29</sup>I used the first wave of surveys from 1994 to 1995 that included in-school data for a sample of 85,627 students in 142 schools. I have information on friends and their SES for 77% of the students surveyed in school. For more information on how the Add Health survey was collected, see Harris et al. (2019).

<sup>30</sup>The Socioeconomic (SES) measure is based on the following variables used to predict household income: Father and Mother's occupation, Father and Mother's education, Father and Mother's employment status, number of individuals in the household, and missing indicators. Refer to Appendix D for a longer discussion of how I define high-SES students using the Add Health data.

<sup>31</sup>Note that students who enroll in extracurricular activities are likely a select group. Among low-SES students with extracurricular activities, the correlation between school composition and friendship formation is slightly stronger at 0.40.

composition and the proportion of high-SES friends is 0.78, and the slope between extracurriculars and the proportion of high-SES friends is 0.71.<sup>32</sup>

Next, I identify how a school's cross-income exposure relates to its students' average likelihood of cross-income friendship as captured by Chetty et al. (2022). I link the Texas administrative data with public high school measures on "economic connectedness" captured by Chetty et al. (2022). Economic connectedness is defined by Chetty et al. (2022) as the proportion of high-SES friends among low-SES individuals on Facebook divided by the share of high-SES individuals in the population.<sup>33</sup> For each high school, I create a measure of exposure with which I capture the average proportion of lower-income students' classmates that are higher income in 2012.<sup>34</sup>

The likelihood that a lower-income student is exposed to a higher-income classmate is highly predictive of a school's economic connectedness, as shown in Figure 11 Panel (a). The correlation between the average proportion of higher-income classmates among lower-income students in a school (exposure) and economic connectedness is between 0.86 and 0.93.<sup>35</sup> The correlation between economic connectedness and exposure to upper-income students is approximately the same if I define exposure as the school average lower-income students' proportion of upper-income classmates or the school's proportion of upper-income students (independent of classroom). The similar correlation is likely because, as demonstrated in Section 4.2, classrooms contribute very little to lower-income students' average exposure to upper-income students. The correlation between economic connectedness and exposure is consistent with Chetty et al. (2022), who find that a 10-percentage-point increase in exposure to higher-income peers in schools increases the likelihood of befriending a higher-income student by 8.9 percentage points.<sup>36</sup>

Chetty et al. (2022) find that, conditional on a school's income composition, students in some schools have a lower likelihood of befriending higher-income students (higher friending bias).<sup>37</sup> Schools' friending bias may capture differences between schools in within-school sorting by income, i.e., conditional on a schools'

---

<sup>32</sup>Low-SES students who enroll in extracurricular activities have four percentage points more high-SES friends than those who do not in the same school and grade. This might reflect the fact that low-SES students with more high-SES friends are more likely to enroll in extracurricular activities or that extracurricular activities provide the opportunity for friendships with high-SES peers—42% of a student's top 5 female and male friends share at least one extracurricular activity.

<sup>33</sup>Since they define high-SES as having above-median income, the share of high-SES friends is divided by 0.5.

<sup>34</sup>I use students enrolled in high school in 2012 instead of 2019 because the friending bias measures are based on Facebook friendship networks for individuals 25–44 as of 2022. I was able to merge 1,132 schools in the Texas data with the high school friendship measures, that is, 51% of schools serving grades 9–12. Of the 1,132 schools, 981 have data on economic connectedness measures.

<sup>35</sup>The range is based on whether I use the economic connectedness measure based on students' own income or students' parental income.

<sup>36</sup>The stronger correlation in the Texas Chetty et al. (2022) data compared to the Add Health data suggests that lower-income students' exposure to upper-income students is better at capturing differences in school average friendship patterns as measured by Chetty et. al (2022) than it is at capturing differences in individual students' close friendship patterns as captured in the Add Health data.

<sup>37</sup>Friending bias is calculated as one minus the share of friends who are high SES divided by the share of individuals in the group who have high SES. If friendships are formed at random (i.e., high- and low-SES individuals have the same likelihood of befriending high-SES individuals), then friending bias would be equal to 0. Refer to Chetty et al. (2022) for more details.

income composition, lower-income students may have different likelihoods of enrolling in the same classroom as higher-income peers. I find that the correlation between within-school sorting by income (as captured by the variance ratio) and friending bias is between 0.61 and 0.67 as shown in Figure 11 Panel (b). This further illustrates how understanding patterns of cross-income classroom exposure is important to understanding cross-income friendship formation.

The patterns of the relationship between exposure to upper-income students and friendship formation suggest that for the average lower-income student, a schools' proportion of upper-income students is highly predictive of her likelihood of forming cross-income friendships.

## 6 How Does Cross-Income Exposure Relate to College Enrollment, Graduation and Employment?

Having established that exposure to upper-income peers relates to friendship formation that has been linked to economic mobility, in this section I examine the relationship between exposure to upper-income peers and lower-income students' access to school resources, college enrollment and employment. I then try to disentangle how much of this relationship is driven by changes in the proportion of higher-achieving students (not changes in peer income) and/or access to school resources.

### 6.1 Sample and Main Outcomes

In this section, I focus on high school exposure to upper-income classmates to allow me to link students' classroom exposure to their college enrollment, graduation, and wages. Students are grouped into cohorts based on the year they are enrolled in grade 9. For example, students enrolled in grade 9 in 2009–2010 are grouped into the 2010 cohort. I observe seven cohorts of students: 2010 to 2016.<sup>38</sup>

When referring to students' baseline test scores, I am referring to students' grade 8 reading test scores, which are standardized within the full sample of students who had taken the grade 8 test in a given year—92% of the students in the sample have grade 8 reading test scores. The data are at the student-school level. Students are assigned to their grade 9 school. Some students are enrolled in multiple schools in grade 9.

---

<sup>38</sup>The income measure is based on the number of years on free/reduced price lunch. I observe each cohort for 10 years. For example, students in the 2010 cohort who never had free/reduced lunch between 2004 and 2013 are identified as higher income.

## 6.2 Descriptive Patterns

Students across test-score groups appear to perform better when exposed to more upper-income classmates during their high school years, as shown in Figure 13. These binned scatter plots are based on lower-income students in cohorts 2011 to 2016 and present the raw observational relationship (with no controls) between classroom exposure to upper-income students in high school and college enrollment and wages. I place students into three groups based on their standardized grade 8 reading test score: students who scored in the top 80th percentile of the cohort test score distribution (11% of lower-income students), students who scored in the bottom 20th percentile (28% of lower-income students), and students who scored between the 20th and 80th percentiles (62% of lower-income students).

Lower-performing students seem to have a stronger relationship between exposure to upper-income students and college enrollment (2- and 4-year) and wages. Lower-performing students are 2.1pp and 1.8pp more likely to enroll in 2- and 4-year colleges, respectively, when in classrooms with 10 pp more upper-income students. They also earn 232 dollars more in early adulthood and have 1.1 higher income percentile rank for every 10pp more upper-income students in their classroom.<sup>39</sup> Top-performing students as measured by test score generally have flatter slopes, with the exception of selective college enrollment and college graduation: they seem to be substantially more likely to enroll in selective universities (1.1pp) and more likely to graduate from college (1.4pp) when exposed to 10 pp more upper-income classmates.

The positive relationship between exposure to upper-income classmates and long-term outcomes could be capturing both systematic differences between lower-income students who attend schools with more upper-income peers and those who attend the same school but classrooms with more upper-income peers. For example, a highly motivated lower-income student may be more likely to find a way to attend a school with more upper-income students and within that school more likely to enroll in advanced courses that happen to have more upper-income peers. In the next section, I address some of the selection concerns by using variation in cohort composition within a school, independent of the classroom a student enrolls in.

The positive relationship between a student's proportion of upper-income classmates and long-term outcomes may be driven by differences in access to school resources correlated with the proportion of upper-income classmates. For example, a lower-income student may be more likely to enroll in college if they attend schools with more upper-income peers because those schools also happen to have more experienced teachers. Across test scores, students exposed to more upper-income students also have fewer novice teachers (3 or less

---

<sup>39</sup>Wages are based on quarterly wages in 2022 for cohorts that were in grade 9 between 2011 and 2014. Income percentiles are based on the distribution of wages in the sample within a given cohort. In Appendix Figure A31, I also present the relationship between exposure to upper-income students and industry of employment. Students exposed to more upper-income students are slightly less likely to enter the construction industry and more likely to enter the professional, scientific and technical services industry.

years of experience) in high school, as shown in Figure 14. Students with 10pp more upper-income classmates are in classrooms with 2.6pp fewer novice teachers, on average.<sup>40</sup> The relationship between advanced course taking and exposure to upper-income students seems more ambiguous, although generally positively correlated with exposure to upper-income students. In comparison, school spending per pupil seems to decrease with higher exposure to upper-income students, as shown in Figure 15 Panel (a): Students with 10pp more upper-income classmates are in schools with 647 dollars lower spending per pupil.<sup>41</sup>

Access to school resources (other than school spending per pupil) may vary both between and within schools. For example, Kalogrides and Loeb (2013) find that lower-income students, compared to upper-income students in the same school, tend to enroll in classrooms with less experienced teachers. Similarly, I find that within a school lower-income students in classrooms with (10pp) more upper-income students tend to have (3pp) fewer novice teachers compared to other lower-income students in the same school. The correlation between teacher experience and the proportion upper-income classmates within a school suggests that even in a school access to school resources may vary with the proportion of upper-income classmates.<sup>42</sup>

### 6.3 Identification Strategy

The relationship between exposure to upper-income peers and long-term outcomes may be driven by systematic differences between students who enroll in districts, schools and classrooms with more upper-income peers. For example, schools with more upper-income peers may also have better teachers that affect students' outcomes. One way to address these selection concerns is to focus on within school variation in the proportion of upper-income peers across cohorts, building on work by Hoxby (2000).<sup>43</sup> The assumption here is that adjacent cohorts of students who attend the same school would have a similar school environment and the only difference between them is that some happen to have more upper-income peers due to random variation between cohorts in peer income composition that is unrelated to changes in school characteristics.

In this section, I use variation in grade 9 school composition for a given cohort, regardless of whether those students end up in the same classroom, unlike in prior sections where I used classroom exposure to upper-income peers. Therefore, the estimates from the regression should be interpreted as the reduced-form effect of variation in school cohort composition on classroom exposure.<sup>44</sup> I use variation in school-cohort

---

<sup>40</sup>Teacher experience seems to matter most in the first few years of experience (see Rice, 2013).

<sup>41</sup>Spending patterns slightly vary across school by proportion upper-income students. Schools with 10 pp more upper-income students spend slightly more of their budget on instruction (0.7pp) and extracurriculars (0.6pp).

<sup>42</sup>Differences in the proportion of upper-income classmates within a school are much more predictive of a lower-income students' number of advanced courses taken than the overall difference in proportion of upper-income classmates.

<sup>43</sup>Since Hoxby (2000) this approach has been used extensively in the peer effects literature. For examples see Feld and Zöllitz (2017), Lavy and Schlosser (2011) and Angrist and Lang (2004).

<sup>44</sup>The reduced-form framing assumes that upper-income students only impact lower-income students if they share the same classroom. This assumption (exclusion restriction) is violated if upper-income students move resources into the school or change

instead of classroom-cohort because I cannot track the same classrooms over time and students who attend different classrooms are likely to be systematically different. I use 7 years of cohort data from 2010 to 2016 and limit the analysis to schools with at least three cohorts of grade 9 students during the period (enough for a time trend). This gives me 2,152 schools and 13,604 school cohorts. For wages and graduation, I only include cohorts from 2010 to 2014 to allow enough time to capture the impact on college graduation and wages.

A concern with using within school variation in cohort composition is that between cohort changes may be related to changes in the school across time. For example, a school might start offering more AP courses and that attracts more upper-income students to the class. To address the time-varying selection concern I include school-time trends. The school time trends capture systematic linear changes in cohort composition that may be driven by school and student unobserved changes.<sup>45</sup> The source of variation that remains is “unexpected shocks” (temporary) deviations from the school time trend in the proportion of upper-income peers. The identifying assumption is that these shocks to cohort composition are effectively random—they only impact students through changing their cohort composition and are not capturing other cohort-specific changes that are correlated with long-term outcomes.

For example, a district may be enrolling more upper-income peers over time captured by a positive linear time-trend in the proportion upper-income students, but on some years there are more/less than predicted proportion upper-income peers due to random variations like an upper-income family having children of age groups that coincide with some cohorts. I show in Section 6.6 that the distribution of deviations from school trend in the proportion of upper-income peers captured here is similar to the distribution of simulated random deviations from school mean.

Using within school cross-cohort deviations from school trends in the proportion of upper-income peers, I estimate the impact of having more upper-income peers in high school on lower-income students’ college enrollment and wages using the following Equation (1):

$$Y_{isc} = \beta_0 + \beta_1 Prophighincome_{-isc} + X'_{isc}\beta_2 + \delta_c + \delta_s + c \times \delta_s + \epsilon_{isc} \quad (1)$$

---

school norms that impact lower-income students through ways other than sharing the same classroom.

<sup>45</sup>School trends may not capture the pattern of school time-varying changes well and so as an additional robustness check in Section 6.6 I assign students their adjacent cohort proportion upper-income peers as a placebo test. I show that adjacent cohorts’ peer income composition do not appear to impact a students’ own outcomes suggesting I am not capturing spurious correlations between the proportion of upper-income peers and time-varying school factors.

where  $Prophighincome_{-isc}$  is the share of student  $i$ 's peers (excluding herself) in cohort  $c$  that are upper income;  $\beta_1$  captures the impact of an unexpected change (residual of the linear school trend) in the composition of peers on the long-term outcome; and  $X'_{isc}$  is a vector of student  $i$ 's characteristics (gender, race/ethnicity and math and reading grade 8 test scores).  $\delta_s$  captures a school fixed effect to account for baseline school differences in student outcomes, and  $c \times \delta_s$  captures school-specific time trends ( $c$  is a linear cohort trend). I also include cohort fixed effects,  $\delta_c$ , to capture any cohort-specific changes in the sample. Because the treatment is at the school-cohort level, I cluster the standard errors at the school level. I run the regression only for lower-income students since I am interested in how exposure to upper-income peers impacts lower-income students' long-term outcomes.<sup>46</sup>

Exposure to upper-income peers is highly correlated with exposure to higher-performing students (correlation of 0.53 among lower-income students). As such, we may be capturing the impact of exposure to high-achieving students, regardless of income. To disentangle the two, I run another regression that controls for peer achievement. In-line with Section 4.3, high-achieving students are defined as those who performed in the top 24 percentiles of their grade 8 reading test score. To capture whether the effect is driven by changes in peer test scores, I control for the proportion of high-performing students in a cohort,  $Prophighachieve_{-isc}$ . If the impact is driven by upper-income students' test scores, the coefficient on upper-income would approach 0 when I control for the proportion of high-achieving peers. The regression I run is shown in Equation (2) below:

$$Y_{isc} = \beta_0 + \beta_1 Prophighincome_{-isc} + \beta_2 Prophighachieve_{-isc} + X'_{isc}\beta_3 + \delta_c + \delta_s + c \times \delta_s + \epsilon_{isc} \quad (2)$$

where  $Prophighachieve_{-isc}$  captures the proportion of high-achieving students in school  $s$  cohort  $c$  based on their standardized grade 8 reading score.

The parameter  $\beta_1$  in Equation (1) captures the impact of small deviations in the proportion of upper-income classmates. A standard deviation in the residual cross-cohort variation in the proportion of upper-income students is 2 percentage points. At the 1st percentile are cohorts with 5.5 percentage points fewer upper-income students, and at the 99th percentile are cohorts with 5.8 percentage points more upper-income students. Table 5 summarizes the residual variation in the composition of the cohort and Figure 16 presents

---

<sup>46</sup>Because it is very computationally intensive to run the regression with school-specific time trends, I ultimately residualized both the dependent variable ( $Y_{isc}$ ) and independent variables ( $Prophighincome_{sc}$ ) on the fixed effects and time trends ( $\delta_c$ ,  $\delta_s$  and  $\delta_{sc}$ ) and then regressed the residuals on one another. It provides me with the same standard errors and coefficients as in Equation (1).

the distribution of the residual variation in cohort composition. The median lower-income student is in a high-school classroom with 17 students, 8% of whom are upper-income. Assuming students are randomly assigned to classrooms, a 2 percentage point change in proportion of upper-income students for the median student is an increase in the number of upper-income classmates from 1.4 to 1.7.<sup>47</sup>

Temporary (small) shocks to cohort composition probably would not change the resources students have access to (e.g., teacher quality, AP courses offered, and school budgets)—consistent with what we see empirically in section 6.5. They also likely would not change classroom norms, which is a potential mechanism through which exposure to upper-income peers may impact long-term outcomes. As such, I am likely capturing the impact of small changes in the number of upper-income peers in a classroom that may impact students through direct interactions like the passage of institutional knowledge or changes of expectations on what colleges may be feasible. It could also be through changes in teacher behavior or expectations in response to changes in classroom income composition.<sup>48</sup>

## 6.4 Results

I find some evidence consistent with the notion that a lower-income student in a cohort with a higher proportion of upper-income peers is more likely to enroll in college and earn higher wages. I find that lower-income students are more likely to enroll in 4-year public colleges when exposed to more upper-income peers, as shown in Tables 8 and 7. The results are based on Equation (1). The coefficient on the change in the upper-income share is 0.038 for lower-income students' 4-year college enrollment, implying that lower-income students' likelihood of enrolling in 4-year public colleges rises by 0.38 percentage points for every 10-percentage-point change in the share of their class that is upper income. I do not find any strong evidence of an impact on a student's likelihood of graduating from college and enrolling in selective colleges. The impact on college enrollment appears to be driven by students who ranked in the bottom 20th percentile of their grade 8 reading tests, as shown in Table 10.<sup>49</sup>

I also find evidence of a positive impact on average wages and income percentile rank.<sup>50</sup> The average quarterly wage of lower-income students by 2022 (by age 22–25) increased by 78 dollars (1.7%), and their

<sup>47</sup>The median lower-income student is in a school-cohort with 582 students and 9% upper-income students, so a 2 percentage point change in proportion of upper-income students for the median student is an increase in the number of upper-income students from 52 to 64.

<sup>48</sup>Using this research design I am not able to identify why having more upper-income peers may have spillovers on lower-income students, but I can provide suggestive evidence on potential mechanisms like peer achievement and changes in access to (observable) school resources.

<sup>49</sup>The estimates are consistent, albeit more statistically significant for college enrollment and graduation outcomes, when I run the regression without school time trends as shown in Tables 8 and 7 Models (1)–(3).

<sup>50</sup>Income percentile rank is based on ranking the full sample of students in a given cohort based on their average quarterly wages in 2022. Students who do not have wage data are assumed to earn 0 in that year

percentile rank (within cohort) increased by 0.33 percentiles for every 10-percentage-point change in the share of upper-income students.<sup>51</sup>

Note that a student's cohort composition may not be representative of her actual exposure to upper-income students. Table 6 presents the impact of having more upper-income peers on lower-income students' proportion of upper-income classmates.<sup>52</sup> I find that a 10-percentage-point increase in the proportion of upper-income peers in one's cohort increases lower-income students' proportion of upper-income classmates by 5.6 percentage points. The diluted impact of peer cohort composition on lower-income students' classroom exposure to upper-income peers may be in part due to within-school sorting by income. I find that a 10-percentage-point increase in the proportion of upper-income students in a cohort increased the level of within-school income sorting between classrooms by 1.3 percentage points, as shown in Table 11. That is, the gap between the proportion of upper-income classmates in upper-income and other students' classrooms increases by 1.3 percentage points.

Some aspect of exposure to upper-income students may be independent of peer test score. To disentangle the impact of peer test scores on income, I control for the proportion of high-achieving students in a given cohort, as shown in Equation (2). If peer income impacts college enrollment only through peer test scores, the coefficient of peer income should approach zero when I control for peer test scores.

Tables 8 and 7, Model (6), present the effect of having more upper-income peers, conditional on having more high-achieving peers, on lower-income students' college enrollment and wages. The estimates in this model are based on Equation (2). I find that the impact of exposure to upper-income peers on college enrollment remains positive and consistent with previous estimates when controlling for peer test score—if anything, it appears to be slightly larger. The impact of exposure to high-achieving students appears to be negative. This suggests that there is something about being exposed to upper-income peers that impacts student college enrollment and wages through mechanisms other than peer achievement (orthogonal to peer test scores).<sup>53</sup>

The negative impact of the proportion higher-achieving students on lower-income students compared to the positive impact of upper-income peers suggests that the effect of upper-income peers is driven by non-high-achieving upper-income peers. In Table A22 I split peers into four categories: upper-income and high-achieving, upper-income and not high-achieving, lower-income and high-achieving and lower-income

---

<sup>51</sup>The percent increase in wages is based on dividing the 78 dollars by the average lower-income students' quarterly wage in the sample in early adulthood (\$4642)

<sup>52</sup>Since I only have classroom data starting in 2011, I limit this analysis to cohorts from 2011 to 2016.

<sup>53</sup>The patterns are consistent independent of how higher-achieving students are defined. I define higher-achieving students as those who scored in the top 24 percentiles of the math (instead of reading) grade 8 test scores in Table A12 and the top 24 percentiles of students' average of math and reading scores in Table A13. In Table A15, I control for a continuous measure of peer test score: peer average reading test scores in grade 8.

and not high-achieving. The patterns of results suggest that the group driving the positive effect on college enrollment and wages is upper-income not high-achieving peers.

The relationship between the proportion of upper-income peers and long-term outcomes may operate through its short-term impact on students' test scores and high school behavioral outcomes (discipline and attendance). As shown in Table 9, I find that a 10-percentage-point increase in the proportion of upper-income peers decreases the probability of lower-income students receiving an out-of-school suspension by 0.43 percentage points and the proportion of days absent by 0.15 percentage points. A 10-percentage-point increase in the proportion of upper-income peers also appears to increase lower-income students' average reading test score by 0.02 standard deviations.<sup>54</sup> In comparison, having more higher-achieving students appears to have no (detectable) impact on students' behavioral outcomes (with the exception of attendance) and a negative impact on lower-income students' test scores.

I attempt to identify the impact of exposure to upper-income students of the same race/ethnicity. Some evidence suggests stronger peer effects within racial groups, but ultimately most estimates are noisy, as shown in Tables A17 to A19. The impact on 4-year public college enrollment is stronger for Black students in cohorts with more upper-income Black students (a 10pp change increases 4-year college enrollment by 4.3pp). The impact on wages also appears to be considerably stronger for White students exposed to more upper-income White students (a 10pp change increases wages and income rank in 2022 by 197 dollars and 1.3 percentile ranks).

The change in the proportion of upper-income students tends to coincide with a change in the cohorts' racial/ethnic composition. It tends to increase the proportion of white students, as shown in Table A11. One might ask if the impact of exposure to upper-income peers is capturing the impact of exposure to more white students. The impact of exposure to upper-income students is consistent with prior estimates conditional on changes in cohort racial/ethnic composition, as shown in Table A16, suggesting that the impact of exposure to upper-income peers is independent of changes in peer racial/ethnic composition.

## 6.5 Addressing Economic Segregation and the Role of Access to School Resources

An increase in the proportion of upper-income peers may impact students' college enrollment by changing students' course-taking patterns and/or access to experienced teachers. If the positive impact of exposure

---

<sup>54</sup>The impact on test scores may be driven by a change in the composition of students taking the test. I do not find evidence that the change in peer income composition impacted the proportion of students missing reading test scores in high school. Most students took the Reading TAKS test in grade 9 (2010-2011) or the English I test in grade 9 (2012-2019).

to upper-income peers is exclusively through access to school resources, then resource equalization policies may be enough to address concerns with economic segregation.

Table 11 presents the impact of having more upper-income peers on lower-income students' access to (observable) school resources: advanced course enrollment, teacher experience and spending per pupil. I find no (detectable) evidence that changes in peer income composition impact students' average teacher experience, number of advanced courses taken or spending per pupil—the coefficients are small and insignificant. For example, a 10 pp increase in the proportion of upper-income peers increases the proportion of novice teachers by 0.001 ( $p = 0.592$ ) and number of AP courses a student takes by 0.02 ( $p = 0.269$ ). However, having upper-income peers may impact lower-income students' access to school resources in a way that we cannot observe in the data. For example, upper-income parents can volunteer to help plan school events or teach elective courses. Since policies can only impact access to school resources observed in the data, unobserved changes in access to school resources would likely require changes in cross-income exposure.

Exposure to upper-income classmates may impact lower-income students via “direct” means through classroom exposure and “indirect” means through moving resources to the district and/or school, regardless of whether students share the same classroom. If the impact is driven by direct exposure, then policies should focus on increasing cross-income classroom exposure. If, instead, it is a matter of indirect exposure (e.g., changes in school budget or courses), then policies could instead focus on how resources are allocated. To capture the impact of “indirect” exposure to upper-income students in the district, I use cross-cohort variation in the proportion of upper-income students in district schools in the same cohort, excluding the student's own school. This “indirect” exposure measure captures the impact of a change in the proportion of upper-income students in schools other than the student's own school in the district.<sup>55</sup>

I do not find any evidence of a district spillover effect on college enrollment and wages or access to school resources. Changes in the income composition of the district do not appear to affect students attending other schools in the same district, as shown in Tables A20 and A21.

The positive (marginal) impact of exposure to upper-income peers on lower-income students suggests that at least some of the gains to improving cross-income exposure do not operate exclusively through changes in access to school resources. As such, resource equalization policies may not be enough to address the harms of economic segregation. To address the gap in exposure to upper-income peers documented in Section 4 we would want to swap upper- and lower-income students to improve lower-income students' exposure to upper-income peers. Since there are three income groups (upper-, middle- and lower-income) I control for

---

<sup>55</sup>Ideally, I would like to differentiate between the impact on long-term outcomes of sharing the same classroom relative to sharing the same school but a different classroom. This is not possible in this setting, since students who attend different classrooms are likely to be systematically different and I cannot follow the same classroom across cohorts.

the proportion of middle-income students to capture the impact of swapping a lower-income student with an upper-income student (and the other way around).

In Tables 12 and 13 I present the impact of changing the proportion lower- and upper-income students, controlling for the proportion of middle-income students in the cohort. By controlling for the proportion of middle-income students I try to simulate an experiment where upper- and lower-income students move around but middle-income students stay since their average exposure to upper-income peers is close to the integration benchmark. The results suggest that a 10 p.p. increase in the proportion upper-income students (holding constant the proportion of middle-income students, i.e. swapping lower-income students with upper-income students) would increase lower-income students' wages in early adulthood by 2.1% ( $p = 0.02$ ). Alternatively, a 10 p.p. increase in the proportion lower-income students (holding constant the proportion of middle-income students) has no detectable impact on upper-income students wages (-0.6% ( $p = 0.5$ )).

The positive impact on lower-income students of increased exposure to upper-income peers and the lack of detectable impact on upper-income students of increased exposure to lower-income peers suggests that improving cross-income exposure could be a win-win/neutral situation. Note that the estimates are based on temporary shocks in exposure to upper-income peers within the same school. As such, it captures the impact on lower- and upper-income students who remain in the same school, but not the impact on upper- and lower-income students who would have to move schools to improve cross-income exposure. It also captures the impact of temporary changes which may be a lower-bound on potential gains from long-term changes in the proportion of upper-income peers on lower-income students if long-term changes are more likely to improve access to school resources. Importantly, these estimates assume that the slope of the relationship between the proportion of upper-income peers and long-term outcomes is linear.

## 6.6 Falsification Test and Common Cohort Shocks

The key identifying assumption is that deviations from the school trend in the proportion of upper-income students are uncorrelated with changes in unobserved factors correlated with lower-income students' outcomes. One way to assess the validity of this assumption is to examine if the deviations from the school trend in the proportion of upper-income peers look like the type of natural fluctuations from the mean proportion of upper-income peers that happen year-to-year. For example, there may be random fluctuations because an upper-income family in the district happen to have children of age groups that match certain cohorts and not others. I use a Monte Carlo Simulation to assess whether the observed within school deviations from school trend in the proportion of upper-income students look like the variation that would result from random draws of upper-income students. For each school, I randomly generate the income of

the students in each cohort using a binomial distribution function with  $p$  equal to the average proportion of upper-income students in the school across all years. I then compute the within school standard deviation of the proportion of upper-income students using residuals from a regression of the proportion upper-income students on school specific time trends. I repeat this process 1,000 times to obtain the empirical 90 percent confidence interval for the standard deviations for each school. Figure 17 shows that the distribution of schools' average simulated standard deviation from school mean is very similar to the observed schools' standard deviation in the proportion of upper-income students from school trend.<sup>56</sup> I find that 88 percent of the schools had a standard deviation within the empirical 90 percent confidence interval of the school's distribution of simulated standard deviations, which is close to what we would expect from random draws of the school population.

To further examine if the estimates are actually capturing spurious correlations between changes in the proportion of upper-income peers and unobserved factors correlated with lower-income students' outcomes in Tables 14 and 15 I present falsification tests based on placebo cohort proportion upper-income peers replacing the true cohort composition with the cohort composition in the younger ( $t - 1$ ) or older cohort ( $t + 1$ ).<sup>57</sup> The results based on using  $t - 1$  and  $t + 1$  placebo cohorts show no detectable effects on any of the outcomes (coefficients are small and insignificant). For example, using cohort  $t + 1$  a 10 p.p. change in proportion upper-income peers is associated with -.015 ( $p = 0.5$ ) change in lower-income students' 4-year college enrollment and -9.4 dollars ( $p = 0.8$ ) change in wages in early adulthood is. The lack of detectable effect suggests that the estimates I find are not driven by spurious correlation between the proportion upper-income students and time-varying school factors and that there are no spillovers across cohorts.

Even if the variation in the proportion of upper-income peers resembles a random process and does not seem to relate to school changes across cohorts, these variations could be correlated with cohort-specific changes (common shocks). For example, lower-income students in cohorts with larger deviations from school trend in the proportion of upper-income students may happen to have higher family income. To assess this possibility, I check whether lower-income students' demographic characteristics are correlated with school deviations from school trend in the proportion of upper-income students. The proportion upper-income students does not seem to correlate with lower-income students' baseline characteristics (gender, race and immigrant status) suggesting that the estimates are not driven by simultaneous changes in lower-income students' characteristics as shown in Table A25. Lower-income students' parental income reported

---

<sup>56</sup>The right tail of the observed distribution of standard deviations is longer than the simulated distribution. This is likely in part because the simulated distribution is based on the average from multiple simulations. I rerun the main estimates excluding schools with standard deviations above 0.3 and the estimates remain consistent as shown in Table A23.

<sup>57</sup>The falsification test is used by Lavy and Schlosser (2011) to identify if the impact of female peers is driven by time-varying school changes.

on financial aid data also appears to be no different when exposed to more upper-income peers, suggesting that the estimates are not driven by cohort specific changes that correlate with better long-term outcomes.<sup>58</sup>

Another set of concerns with this identification strategy relate to how treatment is defined. The deviations in the proportion of upper-income students may be capturing small measurement errors in a students' income, i.e., lower-income students' peer income composition did not change it just happens to be that one student who should have been assigned one year on free/reduced lunch was not assigned to free/reduced lunch due to administrative or other errors. I can check if this is the case by examining how the treatment (proportion of upper-income peers) correlates with a continuous measure of peer income. I find that a 10pp increase in the proportion of upper-income students—defined as students never on free/reduced lunch—is associated with a 5 thousand dollars higher average cohort income based on financial aid applications and 8 pp fewer cohort average proportion of years on free/reduced price lunch status as shown in Table A24.

The treatment—deviations from high school trends in the proportion of upper-income students—might be capturing changes in cohort proportion upper-income peers that date back to earlier school years. In other words, lower-income students in cohorts with more upper-income students in high-school likely also had more upper-income students in middle and elementary schools. I find that a 10pp increase in the proportion of upper-income students in high-school is associated with a 6pp higher proportion of upper-income students in grade 8. The correlation between the proportion upper-income in middle and high school suggests that the treatment may be capturing the cumulative impact of deviations from school trends in the proportion of upper-income students on lower-income students.

To capture the impact of exposure to upper-income students in high school independent of middle school exposure and to address the concern that grade 8 peer test-score may be a function of grade 8 exposure to upper-income peers, I include fixed effects for students' grade 8 school cohort as shown in Table A26 and Table A27 Models (5). By including grade 8 school cohort fixed effects I am limiting the variation to students who attended the same grade 8 school and cohort but happen to attend high schools with differing deviations from school trend in the proportion of upper-income students. The coefficients on the proportion of upper-income peers are generally consistent but insignificant when controlling for grade 8 school cohort, with the exception of public 4-year college enrollment which remains consistent and significant.<sup>59</sup>

---

<sup>58</sup>Note that the financial aid data is based on a select group of students and the proportion of lower-income students who applying for financial aid slightly increases when lower-income students are exposed to more upper-income peers—by 0.5 percentage points for every 10 pp increase in the proportion of upper-income students.

<sup>59</sup>The coefficient on the proportion high-achieving peers is sensitive to model specification. The patterns in Table A27 suggest that the proportion of upper-income students matters independent of students' achievement and model specification, but it is not clear what the impact of exposure to high-achieving students is—it appears to depend on specification, ranging from positive to negative. My preferred specifications are Models (4) and (5) because they include controls for a student's own test score. Models (4) and (5) suggest that the impact of having a higher proportion of high-achieving students is negative on college enrollment and the impact on wages is small and insignificant.

I also present estimates on the impact of peer achievement using earlier student test-scores that may not have been impacted by changes in cohort peer income composition. Ideally, I would want to control for student achievement measured before students enrolls in school, but in the absence of this I use grade 3 test-scores which are the earliest test-scores I have for students. Table A28 shows that the coefficients on the impact of proportion upper-income peers is consistent independent of if I control for grade 3 or grade 8 student peer reading test-scores. That said, the coefficient on the impact of exposure to more high-achieving peers is smaller (less negative) when controlling for grade 3 peer test-scores with the exception of the impact on wages which is similarly negative and significant.

## 6.7 The Role of Friendship Formation

Assume that a student's exposure to upper-income peers has a causal impact on her income in adulthood (as suggested in the results of this paper). If the mechanism requires friendship formation, then the impact of exposure to upper-income students on wages would be smaller than the impact of befriending higher-income students on wages. If instead the impact of exposure to upper-income peers does not require friendship formation and is due to other factors related to exposure, then the impact of exposure to upper-income students may be larger than the impact of befriending higher-income students. Ideally, I would want to compare the causal relationship between friendship and wages to that between exposure and wages. To do so, I would need random variation in each of exposure and friendship. In the absence of this, I examine how the raw (no controls) relationship between friendship and wages compares to the relationship between exposure and wages in the Add Health data. In the Add Health data, I have information on students' reported income by age 24–32.

I find that the correlation between low-SES students' upper-income peers and own income is slightly stronger than the correlation between students' proportion of high-SES friends and own income in early adulthood: correlation 0.07 relative to 0.03. The stronger correlation between exposure and students' income suggests that the relationship between peer composition and long-term outcomes may not require lower-income students to be close friends with upper-income students. Since the Add Health friendship data is based on listing students' top five friends, I cannot eliminate the possibility that the relationship between peer composition and long-term outcomes may require (weak) friendships with upper-income peers. That said, it suggests that information on a lower-income student's exposure to upper-income peers may be more predictive of her wage in adulthood than information on her proportion of close upper-income friends.

I can similarly compare the Chetty et al. (2022) measure of the relationship between economic connectedness and income in adulthood to the relationship between exposure to upper-income students and long-term

outcomes captured in this paper. Using a movers design, Chetty et al. (2022) find that moving to a neighborhood with 1 unit higher economic connectedness increases income in adulthood by 30.7%. The slope of the relationship suggests that a 10 percentage-point difference in a schools' proportion upper-income peers (equivalent to a 0.16 units difference in economic connectedness) would increase income in adulthood by 4.9%, which is much larger than the impact on wages in early adulthood found in this paper (1.7% increase in wages).<sup>60</sup>

To compare the estimates of the impact of economic connectedness in Chetty et al. (2022) to the estimates of the impact of exposure to upper-income peers in this paper, we need to assume that for a lower-income student, the effect of a 10-percentage-point change in the proportion of upper-income peers in the same school is equivalent to moving to a school with a 10-percentage-point difference in the proportion of upper-income peers. I will call this assumption the “movers assumption”. The movers assumption is violated if schools with consistently more upper-income peers benefit from having upper-income students move more resources to the school over the long run (such as more AP courses), while an equivalent temporary shock to peer income composition in a school would not allow for enough time for peer income to change students' access to resources.

Here, the weaker relationship between exposure to upper-income peers and long-term outcomes relative to measures of economic connectedness may be driven by a number of factors. Those include that the effect of a shock to a school's cohort composition is not equivalent to moving to a school with more upper-income peers (i.e., the movers assumption does not hold). If the movers assumption holds, then the smaller effect of exposure to upper-income peers found in this paper could be because economic connectedness is capturing something that is particularly important for long-term outcomes that is not captured by exposure to upper-income students.<sup>61</sup>

## 7 Discussion and Conclusion

I find that that lower-income students are generally exposed to very few upper-income students through grades 5 to 12 in Texas and that students are more clustered by income than they are by achievement. Lower-income students' exposure to upper-income peers is driven mainly by district and school segregation by income. How students are sorted within school classrooms has a much smaller impact on the average

<sup>60</sup>The conversion from economic connectedness units to school proportion of upper-income students is based on the slope of the relationship between a schools' economic connectedness as captured by Chetty et al. (2022) and average lower-income students' proportion of upper-income classmates as captured in Section 5 in this paper.

<sup>61</sup>There are a number of other potential reasons for why the estimates are different which include that in Chetty et al. (2022) the estimate is based on changes in neighborhood and not school economic connectedness. If neighborhood composition has a different effect from school composition, then this could also explain the difference in estimates. It could also be because, in this paper, I capture income in early adulthood, which may be less stable and understates the gain to exposure.

lower-income student's exposure to upper-income peers. Nevertheless, classrooms may be an easier policy lever, and some lower-income students would have gained considerably had students instead been randomly assigned to classrooms, particularly in high school.

Using the residual from school trend cross-cohort variation within schools, I find that having more upper-income peers improves lower-income students' college enrollment and quarterly wages in early adulthood. The effect of peer income does not appear to be driven by differences in peer test scores, suggesting that we should think of peer income spillovers as operating separately (and potentially despite) peer achievement spillovers. Changes in cross-cohort composition do not appear to change students' access to school resources like the proportion of novice teacher or number of AP courses, suggesting that the impact of exposure to upper-income peers captured in this paper is not driven by changes in access to school resources. This implies that policies that aim to equalize access to school resources would not fully address the missed-out-on potential gain from cross-income exposure.

The high level of income segregation documented in this paper and the positive impact of upper-income peers on lower-income students' long-term outcomes suggest that income desegregation policies may be particularly beneficial for economic mobility. Decreasing segregation by income also appears to be a win-win/neutral situation, as I find that upper-income students are not (negatively) impacted by exposure to lower-income peers. If anything, prior work by Rao (2019) finds using a school desegregation policy in India that having lower-income classmates makes upper-income students more prosocial (more likely to volunteer for charitable causes and more giving) and less likely to discriminate against lower-income students.

The estimates of the impact of exposure to upper-income peers are based on small (and temporary) variation in peer income composition. The advantage of using small (temporary) shocks in peer income composition is that these shocks are less likely to impact lower-income students' access to school resources, and so it helps us isolate the effect of peer income spillovers. Nevertheless, the use of temporary shocks also means that the estimates of the impact of peer income might not generalize to policies that result in permanent changes in peer income composition. If we think that permanent changes in the proportion of upper-income peers would also allow for more time for changes in access to school resources like experienced teachers, then the estimates from this paper would underestimate the potential gain from desegregation.

The impact of exposure to upper-income students may be context dependent. Future research should examine how and when peer income matters. For example, it is possible that having upper-income peers matters less when access to information on colleges and employment is more available (e.g., Dynarski, Libassi, Michelmore and Owen, 2021) and/or in industries less dependent on job referrals. It also might matter less if the likelihood of enrolling in certain colleges and working in certain industries is equalized across income

groups (that is, if college enrollment and employment norms do not differ across groups). Future work would benefit from surveys that shed light on the differential availability of information to upper- and lower-income students and the role of classroom composition as it relates to norm setting.

## 8 References

- Alcaino, M., & Jennings, Jennifer. L. (2020). How Increased School Choice Affects Public School Enrollment and School Segregation. <https://doi.org/10.26300/83XJ-8E66>
- Angrist, J. D., & Lang, K. (2004). Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *American Economic Review*, 94(5), 1613–1634. <https://doi.org/10.1257/0002828043052169>
- Antonovics, K., Black, S. E., Cullen, J. B., & Meiselman, A. Y. (2022). Patterns, Determinants, and Consequences of Ability Tracking: Evidence from Texas Public Schools (Working Paper No. 30370). National Bureau of Economic Research. <https://doi.org/10.3386/w30370>
- Avery, C., & Pathak, P. A. (2021). The Distributional Consequences of Public School Choice. *American Economic Review*, 111(1), 129152. <https://doi.org/10.1257/aer.20151147>
- Bergman, Peter (2018). The Risks and Benefits of School Integration for Participating Students: Evidence from a Randomized Desegregation Program, IZA Discussion Papers, No. 11602, Institute of Labor Economics (IZA), Bonn
- Biasi, B. (2019). School Finance Equalization Increases Intergenerational Mobility: Evidence from a Simulated-Instruments Approach (Working Paper 25600). National Bureau of Economic Research. <https://doi.org/10.3386/w25600>
- Bischoff, K., & Owens, A. (2019). The Segregation of Opportunity: Social and Financial Resources in the Educational Contexts of Lower- and Higher-Income Children, 1990–2014. *Demography*, 56(5), 1635–1664. <https://doi.org/10.1007/s13524-019-00817-y>
- Billings, S. B., Deming, D. J., & Rockoff, J. (2014). School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg\*. *The Quarterly Journal of Economics*, 129(1), 435476. <https://doi.org/10.1093/qje/qjt026>
- Black, S. E., Devereux, P. J., & Salvanes, K. G. (2013). Under Pressure? The Effect of Peers on Outcomes of Young Adults. *Journal of Labor Economics*, 31(1), 119–153. <https://doi.org/10.1086/666872>
- Bourdieu P (1997) The forms of capital. In:Halsey AH, Lauder H, Brown P, et al. (eds) Education: Culture, Economy, Society, pp. 4658. Oxford: Oxford University Press.
- Bursztyn, L., & Jensen, R. (2015). How Does Peer Pressure Affect Educational Investments? *The Quarterly Journal of Economics*, 130(3), 13291367. <https://doi.org/10.1093/qje/qjv021>
- Bursztyn, L., Gonzalez, A. L., & Yanagizawa-Drott, D. (2020). Misperceived Social Norms: Women Work-

- ing Outside the Home in Saudi Arabia. *American Economic Review*, 110(10), 29973029. <https://doi.org/10.1257/aer.20180975>
- Campbell, Jordan & Smith, Aaron Garth. (2021) Analysis of Texas School District Open Enrollment Data, Reason Foundation. [https://reason.shinyapps.io/texas\\_student\\_transfer\\_dashboard/](https://reason.shinyapps.io/texas_student_transfer_dashboard/)
- Card, D., & Giuliano, L. (2016). Universal screening increases the representation of low-income and minority students in gifted education. *Proceedings of the National Academy of Sciences*, 113(48), 1367813683. <https://doi.org/10.1073/pnas.1605043113>
- Cattan, S., Salvanes, K. G., & Tominey, E. (2022). First Generation Elite: The Role of School Networks.
- Chetty, R., Grusky, D., Hell, M., Hendren, N., Manduca, R., & Narang, J. (2017). The fading American dream: Trends in absolute income mobility since 1940. 9.
- Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenburg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., Wernerfelt, N. (2022a). Social capital I: Measurement and associations with economic mobility. *Nature*, 608(7921), Article 7921. <https://doi.org/10.1038/s41586-022-04996-4>
- Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenburg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., Wernerfelt, N. (2022b). Social capital II: Determinants of economic connectedness. *Nature*, 608(7921), Article 7921. <https://doi.org/10.1038/s41586-022-04997-3>
- Clotfelter, C. T., Ladd, H. F., Clifton, C. R., & Turaeva, M. R. (2021). School Segregation at the Classroom Level in a Southern New Destination State. *Race and Social Problems*, 13(2), 131160. <https://doi.org/10.1007/s12552-020-09309-w>
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2002). Segregation and Resegregation in North Carolinas Public School Classrooms Do Southern Schools Face Rapid Resegregation. *North Carolina Law Review*, 81(4), 14631512.
- Cohodes, S. R. (2020). The Long-Run Impacts of Specialized Programming for High-Achieving Students. *American Economic Journal: Economic Policy*, 12(1), 127166. <https://doi.org/10.1257/pol.20180315>
- Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, 94, S95S120.
- Collins, C., & Gan, L. (2013). Does Sorting Students Improve Scores? An Analysis of Class Composition (w18848; p. w18848). National Bureau of Economic Research. <https://doi.org/10.3386/w18848>

- Conger, D., Long, M. C., & Raymond, M. J. (2020). Advanced Placement and Initial College Enrollment: Evidence from an Experiment. <https://doi.org/10.26300/CX24-VX18>
- Cools, A., Fernández, R., & Patacchini, E. (2019). Girls, Boys, and High Achievers. NBER Working Paper Series.
- Currarini, S., Jackson, M. O., & Pin, P. (2009). An Economic Model of Friendship: Homophily, Minorities, and Segregation. *Econometrica*, 77(4), 1003–1045. <https://doi.org/10.3982/ECTA7528>
- Marcotte, D. E., & Dalane, K. (2019). Socioeconomic Segregation and School Choice in American Public Schools. *Educational Researcher*, 48(8), 493–503. <https://doi.org/10.3102/0013189X19879714>
- Dalane, K., & Marcotte, D. E. (2022). The Segregation of Students by Income in Public Schools. *Educational Researcher*, 51(4), 245254. <https://doi.org/10.3102/0013189X221081853>
- De Giorgi, G., Pellizzari, M., & Redaelli, S. (2010). Identification of Social Interactions through Partially Overlapping Peer Groups. *American Economic Journal: Applied Economics*, 2(2), 241–275.
- Domina, T., Penner, A., & Penner, E. (2017). Categorical Inequality: Schools As Sorting Machines. *Annual Review of Sociology*, 43, 311330. <https://doi.org/10.1146/annurev-soc-060116-053354>
- Duflo, E., Dupas, P., & Kremer, M. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review*, 101(5), 17391774. <https://doi.org/10.1257/aer.101.5.1739>
- Dynarski, S., Libassi, C. J., Michelmore, K., & Owen, S. (2021). Closing the Gap: The Effect of Reducing Complexity and Uncertainty in College Pricing on the Choices of Low-Income Students. *American Economic Review*, 111(6), 1721–1756. <https://doi.org/10.1257/aer.20200451>
- Epple, D., Newlon, E., & Romano, R. (2002). Ability tracking, school competition, and the distribution of educational benefits. *Journal of Public Economics*, 83(1), 148. [https://doi.org/10.1016/S0047-2727\(00\)00175-4](https://doi.org/10.1016/S0047-2727(00)00175-4)
- Feld, J., & Zölitz, U. (2017). Understanding Peer Effects: On the Nature, Estimation, and Channels of Peer Effects. *Journal of Labor Economics*, 35(2), 387–428.
- Fernandez, R. M., & Fernandez-Mateo, I. (2006). Networks, Race, and Hiring. *American Sociological Review*, 71(1), 42–71. <https://doi.org/10.1177/000312240607100103>
- Figlio, D. N., & Page, M. E. (2002). School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality? *Journal of Urban Economics*, 51(3), 497514. <https://doi.org/10.1006/juec.2001.2255>

- Figlio, D. N., & Āzek, U. (2023). The Unintended Consequences of Test-Based Remediation (Working Paper 30831). National Bureau of Economic Research. <https://doi.org/10.3386/w30831>
- Frey, B. S., & Meier, S. (2004). Social Comparisons and Pro-social Behavior: Testing Conditional Cooperation in a Field Experiment. *American Economic Review*, 94(5), 17171722. <https://doi.org/10.1257/0002828043052187>
- Gerring, J. (2012). Mere Description. *British Journal of Political Science*, 42(4), 721746.
- Harris, K.M., C.T. Halpern, E.A. Whitsel, J.M. Hussey, L. Killeya-Jones, J. Tabor, & S.C. Dean. (2019). Cohort Profile: The National Longitudinal Study of Adolescent to Adult Health (Add Health). *International Journal of Epidemiology* 48(5):1415-1425 <https://doi.org/10.1093/ije/dyz115>.
- Hoxby, C. (2000). Peer Effects in the Classroom: Learning from Gender and Race Variation (Working Paper 7867; Working Paper Series). National Bureau of Economic Research. <https://doi.org/10.3386/w7867>
- Hoxby, C. M., & Weingarth, G. (2005). Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects. Working Paper.
- Imberman, S. A., Kugler, A. D., & Sacerdote, B. I. (2012). Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees. *American Economic Review*, 102(5), 2048–2082. <https://doi.org/10.1257/aer.102.5.2048>
- Jackson, C. K. (2013). Can higher-achieving peers explain the benefits to attending selective schools? Evidence from Trinidad and Tobago. *Journal of Public Economics*, 108, 63–77. <https://doi.org/10.1016/j.jpubeco.2013.09.007>
- Jackson, C. K. (2014). Do College-Preparatory Programs Improve Long-Term Outcomes? *Economic Inquiry*, 52(1), 7299. <https://doi.org/10.1111/ecin.12040>
- Klopfenstein, K. (2004). Advanced Placement: Do minorities have equal opportunity? *Economics of Education Review*, 23(2), 115131. [https://doi.org/10.1016/S0272-7757\(03\)00076-1](https://doi.org/10.1016/S0272-7757(03)00076-1)
- Kalogrides, D., & Loeb, S. (2013). Different Teachers, Different Peers: The Magnitude of Student Sorting Within Schools. *Educational Researcher*, 42(6), 304-316. <https://doi.org/10.3102/0013189X13495087>
- Kerckhoff, A. C. (1995). Institutional Arrangements and Stratification Processes in Industrial Societies. *Annual Review of Sociology*, 21(1), 323347. <https://doi.org/10.1146/annurev.so.21.080195.001543>
- Lavy, V., Paserman, M. D., & Schlosser, A. (2012). Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers In the Classroom. *The Economic Journal*, 122(559), 208–237. <https://doi.org/10.1111/j.1468-0297.2011.02463.x>
- Loeb, S., Dynarski, S., McFarland, D., Morris, P., Reardon, S., & Reber, S. (2017). Descriptive Analysis in Education: A Guide for Researchers. NCEE 2017-4023. In National Center for Education Evaluation and Re-

gional Assistance. National Center for Education Evaluation and Regional Assistance. <https://eric.ed.gov/?id=ED573325>

Lucas, S. R., & Berends, M. (2002). Sociodemographic Diversity, Correlated Achievement, and De Facto Tracking. *Sociology of Education*, 75(4), 328348. <https://doi.org/10.2307/3090282>

Mallah, Farah (2024). Impact of Advanced Courses on Cross-Income Exposure and Long-term Outcomes. Working Paper.

Mele, A. (2020). Does School Desegregation Promote Diverse Interactions? An Equilibrium Model of Segregation within Schools. *American Economic Journal: Economic Policy*, 12(2), 228–257.

Michelman, V., Price, J., & Zimmerman, S. D. (2022). Old Boys Clubs and Upward Mobility Among the Educational Elite. *The Quarterly Journal of Economics*, 137(2), 845909. <https://doi.org/10.1093/qje/qjab047>

Moody (2001). Race, School Integration, and Friendship Segregation in America. *American Journal of Sociology*. 107(3), 679-716. <https://doi.org/10.1086/338954>

Monarrez, T., Kisida, B., & Chingos, M. (2022). The Effect of Charter Schools on School Segregation. *American Economic Journal: Economic Policy*, 14(1), 301340. <https://doi.org/10.1257/pol.20190682>

Nechyba, T. J. (2006). Chapter 22 Income and Peer Quality Sorting in Public and Private Schools. In E. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education* (Vol. 2, pp. 13271368). Elsevier. [https://doi.org/10.1016/S1574-0692\(06\)02022-8](https://doi.org/10.1016/S1574-0692(06)02022-8)

Owens, A., Reardon, S. F., & Jencks, C. (2016). Income Segregation Between Schools and School Districts. *American Educational Research Journal*, 53(4), 11591197. <https://doi.org/10.3102/0002831216652722>

Owens, A., & Candipan, J. (2019). Social and spatial inequalities of educational opportunity: A portrait of schools serving high- and low-income neighbourhoods in US metropolitan areas. *Urban Studies*, 56(15), 31783197. <https://doi.org/10.1177/0042098018815049>

Phillips, K. J. R., Larsen, E. S., & Hausman, C. (2015). School choice & social stratification: How intradistrict transfers shift the racial/ethnic and economic composition of schools. *Social Science Research*, 51, 3050. <https://doi.org/10.1016/j.ssresearch.2014.12.005>

Rao, G. (2019). Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools. *American Economic Review*, 109(3), 774809. <https://doi.org/10.1257/aer.20180044>

Rice, J. K. (2013). Learning from Experience? Evidence on the Impact and Distribution of Teacher Experience and the Implications for Teacher Policy. *Education Finance and Policy*, 8(3), 332–348. [https://doi.org/10.1162/EDFP\\_a\\_00010](https://doi.org/10.1162/EDFP_a_00010)

Roy, Susha. (2022) Impacts of Public School Choice on Neighborhoods: Evidence from Los Angeles

Impacts of School. Job Market Paper. Retrieved September 21, 2023

Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates\*. *The Quarterly Journal of Economics*, 116(2), 681704. <https://doi.org/10.1162/00335530151144131>

Sacerdote, B. (2011). Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? In *Handbook of the Economics of Education* (Vol. 3, pp. 249277). Elsevier. <https://doi.org/10.1016/B978-0-444-53429-3.00004-1>

Schmutte, I. M. (2015). Job Referral Networks and the Determination of Earnings in Local Labor Markets. *Journal of Labor Economics*, 33(1), 1–32. <https://doi.org/10.1086/677389>

Setren, E. (2024). Busing to Opportunity? The Impacts of the METCO Voluntary School Desegregation Program on Urban Students of Color.

Zimmerman, D. J. (2003). Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. *The Review of Economics and Statistics*, 85(1), 923.

Zimmerman, S. D. (2019). Elite Colleges and Upward Mobility to Top Jobs and Top Incomes. *American Economic Review*, 109(1), 147. <https://doi.org/10.1257/aer.20171019>

## 9 Main Tables and Figures

Table 1: Texas State Public School Student Demographics: Cohorts 2012-2014

Variable	Mean
Asian-American	.048 (.213)
Black Students	.141 (.348)
White Students	.320 (.466)
Hispanic Students	.516 (.500)
Enrolled in Title 1 School	.572 (.495)
Free/Reduced Lunch Status	.608 (.488)
Special Education	.094 (.292)
Gifted Program	.106 (.307)
Enrolled in Bilingual Program	.027 (.161)
English Language Learner	.262 (.440)
Any Vocational Ed	.353 (.478)
Number of Students	1128554

*Notes.* This table summarizes demographic characteristics of the three cohorts of students I follow from grade 5 to expected grade 12. The demographic characteristics are based on student demographics (including free/reduced lunch status in fifth grade). The average is weighted by number of students across three cohorts. A cohort is defined by the year I observe them enrolled in grade 5 (2012, 2013 or 2014). The number in the brackets is the standard deviation.

Table 2: The Role of Academic Preparation in Accounting for Within-School Differences in Exposure to Upper-Income Students

	(1)	(2)	(3)	(4)
Upper-Income	0.115 (0.0018)	0.0997 (0.0016)	0.0935 (0.0013)	0.0927 (0.0013)
School Fixed-Effect	X	X	X	X
Cohort-Year Fixed-Effect	X	X	X	X
Grade 8 Test-Scores		X	X	X
Algebra I by G8		X	X	X
Attendance and Suspension G8				X
N	4207473	4207473	4207473	4207473
R2	0.916	0.926	0.931	0.931

The upper-income variable is an indicator variable that takes value 1 if a student is upper income. Model 1 includes school and cohortxyear fixed effects. Model 2 additionally includes test-score controls: grade 8 reading and math test scores as well as indicators for missing a test-score. Model 3 additionally includes an indicator for whether a student has taken Algebra I by grade 8. The sample is limited to students in cohorts 2012 to 2014 (based on grade 5 entry) high-school enrollment years. Standard errors are clustered at the school level. The attendance and suspension variables are the following: number of days absent, number of member days, total suspensions (in- and out-of-school) and a missing attendance indicator. Attendance and suspension data are based on expected grade 8 school years.

Table 3: Number and Size of 9th Grade and Demographics of 9th Graders in Texas

Cohort	N. Schools	Median Size	High-Income	High-Achieve	Black	Hispanic	White
2010	1981	68	.276	.295	.14	.472	.314
2011	2067	65	.275	.277	.138	.481	.31
2012	2245	65	.26	.274	.137	.473	.293
2013	2254	68	.255	.247	.137	.481	.286
2014	2266	68.5	.255	.23	.136	.483	.288
2015	2297	71	.253	.237	.135	.491	.284
2016	2275	73	.243	.241	.136	.499	.275

*Notes.* This table summarizes characteristics of the 2010 to 2016 9th grade cohorts. The median school size is based on equally weighting each school. The demographics are student averages weighted by the number of students enrolled.

Table 4: Grade 8 Standardized Reading Scores for 9th Graders

Cohort	% Missing	All	High-Income	Low-Income	High-Achieve
2010	7.4	.017	.384	-.246	.686
2011	8	.024	.409	-.246	.723
2012	6.8	-.011	.409	-.287	.734
2013	7.1	.082	.576	-.201	1.003
2014	8.5	-.026	.172	-.159	.907
2015	7.4	-.088	.098	-.214	.943
2016	7.6	-.198	.139	-.403	1.044

*Notes.* This table summarizes test scores of the 2010 to 2016 9th grade cohorts. The test scores are based on the grade 8 reading test and are standardized within the full population of students who took the test in a given year. “High-Achieve” students are those who scored in the top 24 percentiles of their cohort distribution of standardized test scores.

Table 5: Variation of interest: Between Cohorts 2010 and 2016–Residual Change in Income and Achievement of 9th Graders

Statistic	high-income	high-achieve	high-incomexblack	high-incomexhis	high-incomexwhite
Standard Deviation	.02	.033	.007	.01	.029
1st Percentile	-.055	-.094	-.018	-.027	-.079
5th Percentile	-.027	-.047	-.01	-.014	-.043
10th Percentile	-.018	-.031	-.006	-.009	-.028
90th Percentile	.018	.032	.007	.009	.027
95th Percentile	.028	.048	.01	.015	.045
99th Percentile	.058	.094	.018	.031	.092

*Notes.* This table summarizes the residual variation in the proportion of high-income students, the proportion of high-achieving students, the proportion of high-income and Black students, the proportion of high-income and Hispanic students and the proportion of high-income and White students. The residual variation is based on Equation (1): it is the residual variation after including cohort and school-trend controls. It is based on the sample of low-income students, i.e., it is weighted by the number of low-income students enrolled, which is the population of interest here.

Table 6: Impact of an Increase in the Proportion of Higher-Income Students on Classroom Proportion Upper-Income Classmates for Lower-Income Students

	(1)	(2)
Proportion Upper-Income	0.564 (0.00861)	0.578 (0.00818)
Proportion Upper-IncomeXTop G8 Test Score		0.126 (0.00513)
Proportion Upper-IncomeXLow G8 Test Score		-0.0702 (0.00336)
N Clusters	2103	2094
N	882034	805717

Standard errors are in parentheses and clustered at the school level. Coefficients are based on Equation (1). The regressions include the full sample of lower-income students. The second model includes an interaction with students who scored in the top 80th percentile and bottom 20th percentile based on their grade 8 reading test scores. Models include cohorts 2011–2016 because classroom data are only available starting 2011. The outcome is a student’s classroom proportion of high-income peers. The regression is based on lower-income students.

Table 7: Impact of Peer Composition on College Enrollment for Lower-Income Students

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
<i>Enrolled in College/University</i>						
Proportion Upper-Income	0.0795 (0.0240)		0.108 (0.0246)	0.0455 (0.0246)		0.0640 (0.0249)
Proportion High-Achieve		-0.0869 (0.0210)	-0.0986 (0.0212)		-0.0802 (0.0207)	-0.0854 (0.0208)
Group Mean	0.380					
<i>Enrolled in a Public 2yr College</i>						
Proportion Upper-Income	0.0854 (0.0251)		0.105 (0.0263)	0.0374 (0.0244)		0.0517 (0.0253)
Proportion High-Achieve		-0.0568 (0.0198)	-0.0682 (0.0203)		-0.0619 (0.0201)	-0.0661 (0.0205)
Group Mean	0.290					
<i>Enrolled in a Public 4yr College</i>						
Proportion Upper-Income	0.0411 (0.0165)		0.0542 (0.0163)	0.0379 (0.0193)		0.0479 (0.0194)
Proportion High-Achieve		-0.0401 (0.0148)	-0.0460 (0.0148)		-0.0422 (0.0159)	-0.0461 (0.0159)
Group Mean	0.140					
<i>Enrolled in a Selective University</i>						
Proportion Upper-Income	0.00557 (0.00465)		0.00892 (0.00474)	0.00305 (0.00580)		0.00568 (0.00590)
Proportion High-Achieve		-0.0108 (0.00397)	-0.0118 (0.00402)		-0.0117 (0.00454)	-0.0122 (0.00460)
Group Mean	0.0200					
N	1013905	1013905	1013905	1013905	1013905	1013905
N Clusters	2141	2141	2141	2141	2141	2141
Linear School Time Trends				X	X	X

Standard errors are in parentheses and clustered at the school level. Coefficients are based on Equation (1). Models (1)–(3) do not include school time trends. College enrollment outcomes include cohorts 2010–2016. Models (3) and (6) include both the proportion of upper-income and higher-achieving peers in the cohort. Group Mean is the average for lower-income students in the sample.

Table 8: Impact of Peer Composition on College Graduation and Wages for Lower-Income Students

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
<i>Graduated College/University</i>						
Proportion Upper-Income	0.0552 (0.0206)		0.0589 (0.0206)	0.0272 (0.0226)		0.0308 (0.0227)
Proportion High-Achieve		-0.00534 (0.0175)	-0.0125 (0.0176)		-0.0124 (0.0173)	-0.0154 (0.0174)
Group Mean	0.160					
<i>Average Quarterly Wage 2022</i>						
Proportion Upper-Income	614.8 (323.7)		678.8 (326.5)	778.5 (398.4)		894.1 (401.5)
Proportion High-Achieve		-139.4 (223.4)	-222.6 (224.7)		-415.6 (250.7)	-502.8 (251.9)
Group Mean	4642.8					
<i>Income Percentile 2022</i>						
Proportion Upper-Income	1.399 (1.380)		2.184 (1.401)	3.300 (1.648)		3.745 (1.664)
Proportion High-Achieve		-2.462 (1.019)	-2.730 (1.028)		-1.572 (1.075)	-1.938 (1.080)
Group Mean	46.80					
N	700245	700245	700245	700245	700245	700245
N Clusters	1992	1992	1992	1992	1992	1992
Linear School Time Trends			X	X	X	X

Standard errors are in parentheses and clustered at the school level. Coefficients are based on Equation (1). Models (3) and (6) include both the proportion of upper-income and higher-achieving peers in the cohort. College graduation and wages include cohorts 2010–2014. Group Mean is the average for lower-income students in the sample.

Table 9: Impact of Peer Composition on High School Academic and Behavioral Outcomes

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
<i>Reading/English I Test Score</i>						
Proportion Upper-Income	0.340 (0.0758)		0.476 (0.0788)	0.185 (0.0592)		0.293 (0.0621)
Proportion High-Achieve		-0.422 (0.0681)	-0.474 (0.0684)		-0.464 (0.0651)	-0.488 (0.0652)
Group Mean	-0.300					
N	948421					
<i>Missing Reading/English I Test Score</i>						
Proportion Upper-Income	0.00527 (0.0179)		0.0210 (0.0187)	0.0123 (0.0255)		0.0154 (0.0258)
Proportion High-Achieve		-0.0530 (0.0175)	-0.0553 (0.0180)		-0.0127 (0.0230)	-0.0140 (0.0232)
Group Mean	0.0600					
N	1013905					
<i>Proportion of Days Absent</i>						
Proportion Upper-Income	-0.0200 (0.00703)		-0.0158 (0.00695)	-0.0152 (0.00556)		-0.0136 (0.00557)
Proportion High-Achieve		-0.0166 (0.00484)	-0.0149 (0.00483)		-0.00839 (0.00448)	-0.00726 (0.00449)
Group Mean	0.0900					
N	979552					
<i>Any Out-of-School Suspension</i>						
Proportion Upper-Income	-0.00383 (0.0260)		-0.00728 (0.0265)	-0.0426 (0.0199)		-0.0437 (0.0200)
Proportion High-Achieve		0.0114 (0.0174)	0.0121 (0.0177)		0.00139 (0.0136)	0.00498 (0.0136)
Group Mean	0.210					
N	1013905					
<i>Any In-School Suspension</i>						
Proportion Upper-Income	-0.0436 (0.0424)		-0.0457 (0.0425)	-0.0533 (0.0280)		-0.0499 (0.0281)
Proportion High-Achieve		0.00255 (0.0268)	0.00753 (0.0269)		-0.0200 (0.0195)	-0.0160 (0.0195)
Group Mean	0.360					
N	1013905					

Standard errors are in parentheses and clustered at the school level. Coefficients are based on Equation (1). Group Mean is the average for lower-income students in the sample. All outcome regressions are based on the 2010–2016 cohorts. The test score outcome has fewer observations since 6% of students are missing their reading scores. Most students take the Reading/English I test in grade 9. For students who took the test in 2010 and 2011, their scores were based on the Reading TAKS test. For students who took the test after 2011, their scores were based on the English I STAAR test. All raw scores were standardized based on the distribution of students who took the test that year.

Table 10: Impact of an Increase in the Proportion of Higher-Income Students on Lower-Income Students: By Student Test Score

	(1) Any College	(2) 2yr Public	(3) 4yr Public	(4) Selective	(5) Grad Any	(6) Wage 2022	(7) Income Percentile
Proportion Upper-Income	0.0288 (0.0262)	0.0366 (0.0264)	0.0319 (0.0203)	0.000608 (0.00627)	-0.000156 (0.00843)	799.2 (391.4)	3.883 (1.721)
Proportion Upper-IncomeXTop G8 Test Score	-0.000384 (0.0153)	0.00561 (0.0198)	-0.0829 (0.0170)	0.00506 (0.00849)	0.00736 (0.00913)	-28.98 (193.9)	-1.618 (0.814)
Proportion Upper-IncomeXLow G8 Test Score	0.0558 (0.0163)	0.0141 (0.0135)	0.0452 (0.0137)	0.00370 (0.00243)	-0.000569 (0.00238)	-40.11 (131.7)	1.603 (0.631)
N Clusters	2132	2132	2132	2132	1989	1989	1989
N	925930	925930	925930	925930	638590	638590	638590

Standard errors are in parentheses and clustered at the school level. Coefficients are based on Equation (1) with student own test score interactions. The models include an interaction with students who scored in the top 80th percentile and bottom 20th percentile on their grade 8 test scores. The regression is limited to lower-income students with grade 8 test-scores. Models (1)–(4) include cohorts 2010–2016. Models (5)–(7) include cohorts 2010–2014.

Table 11: Impact of an Increase in the Proportion of Higher-Income Students on Access to School Resources

	(1) Class Sorting	(2) Teacher Experience	(3) Novice Teachers	(4) Spending	(5) Advanced Courses	(6) AP Courses
Proportion Upper-Income	0.129 (0.0138)	0.296 (0.403)	0.0121 (0.0226)	-585.7 (483.0)	0.143 (0.307)	0.232 (0.210)
Control Mean	0.089	10.610	0.281	9835.850	2.558	1.423
N Clusters	1965	1982	1982	1749	2141	2141
N	841215	736994	736994	707294	1013905	1013905

Standard errors are in parentheses and clustered at the school level. Coefficients are based on Equation (1). Model (1) includes cohorts 2011–2016 because the classroom data needed to calculate the within-school variance ratio are only available starting in 2011. Models (2)–(4) include cohorts 2012–2016 because teacher and school budget data linked to classrooms are only available starting in 2012. Spending is average school spending per pupil in a given year. Models (5)–(6) include cohorts 2010–2016.

Table 12: Impact of Proportion Lower-Income Students, Holding Constant Mid-Income Students: on Higher-Income Students

	(1) Any College	(2) 2yr Public	(3) 4yr Public	(4) Selective	(5) Grad Any	(6) Wage 2022	(7) Income Percentile
Proportion Lower-Income	-0.0354 (0.0304)	0.00269 (0.0325)	-0.0432 (0.0330)	0.00184 (0.0207)	-0.0727 (0.0416)	-456.2 (669.9)	-1.261 (2.296)
Proportion Middle-Income	-0.0639 (0.0251)	-0.0213 (0.0277)	-0.0744 (0.0274)	-0.00540 (0.0184)	-0.0883 (0.0338)	-650.8 (557.2)	-1.935 (1.909)
N_clust	1945	1945	1945	1945	1818	1818	1818
N	771243	771243	771243	771243	549280	549280	549280

Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (1) but using variation in the proportion of lower-income students and limiting the sample to upper-income students. It also additionally controls for the proportion of middle-income students. Models (1)–(4) include cohorts 2010–2016 and Models (5)–(7) include cohorts 2010–2014.

Table 13: Impact of Proportion of Upper-Income, Holding Constant Proportion Mid-Income Students: on Lower-Income Students

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Any College	2yr Public	4yr Public	Selective	Grad Any	Wage 2022	Income Percentile
Proportion Upper-Income	0.0240 (0.0268)	0.0256 (0.0266)	0.00887 (0.0216)	0.00244 (0.00634)	0.0236 (0.0253)	984.9 (416.5)	4.822 (1.760)
Proportion Middle-Income	-0.0403 (0.0189)	-0.0222 (0.0191)	-0.0545 (0.0153)	-0.00114 (0.00408)	-0.00662 (0.0178)	374.2 (243.2)	2.759 (1.115)
N_clust	2141	2141	2141	2141	1992	1992	1992
N	1013905	1013905	1013905	1013905	700245	700245	700245

Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (1) but controlling for the proportion of middle-income students. Models (1)–(4) include cohorts 2010–2016 and Models (5)–(7) include cohorts 2010–2014.

Table 14: Impact of Increase in Proportion Upper-Income Students (t+1)

	(1) Any College	(2) 2yr Public	(3) 4yr Public	(4) Selective	(5) Grad Any	(6) Wage 2022	(7) Income Percentile
Proportion Upper-Income (t+1)	-0.00731 (0.0282)	-0.0108 (0.0274)	-0.0151 (0.0213)	0.00249 (0.00637)	-0.00720 (0.0245)	-93.46 (436.6)	0.337 (1.763)
N_clust	2034	2034	2034	2034	1967	1967	1967
N	844665	844665	844665	844665	695382	695382	695382

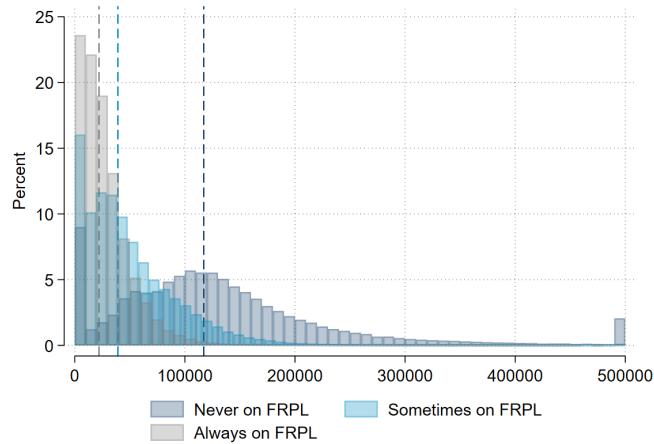
Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (1) but using variation in prior to own cohort (t+1) proportion of lower-income students. Because I use students' t+1 cohort composition, all models do not include 2016 cohorts. Models (1)–(4) include cohorts 2010–2015 and Models (5)–(7) include cohorts 2010–2014.

Table 15: Impact of Increase in Proportion Upper-Income Students (t-1)

	(1) Any College	(2) 2yr Public	(3) 4yr Public	(4) Selective	(5) Grad Any	(6) Wage 2022	(7) Income Percentile
Proportion Upper-Income (t-1)	0.000225 (0.0271)	-0.0261 (0.0256)	0.0325 (0.0217)	-0.00401 (0.00712)	-0.00125 (0.0287)	-430.6 (454.4)	-1.448 (2.017)
N_clust	2076	2076	2076	2076	1929	1929	1929
N	868245	868245	868245	868245	560867	560867	560867

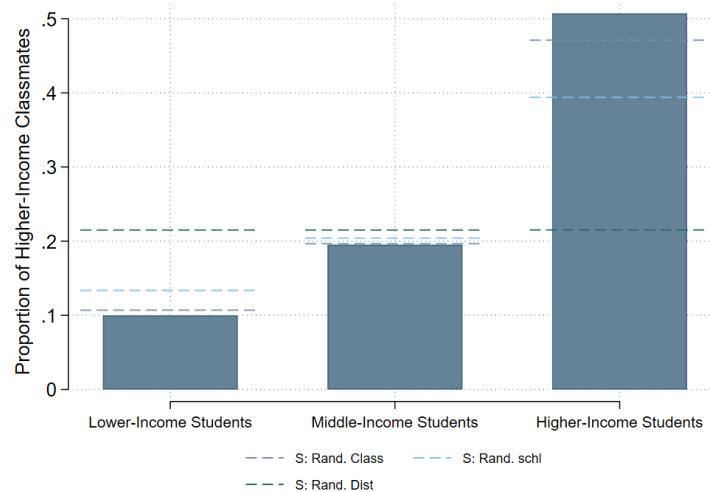
Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (1) but using variation in lagged cohort (t-1) proportion of lower-income students. Because I use students' t-1 cohort composition, all models do not include 2010 cohorts. Models (1)–(4) include cohorts 2011–2016 and Models (5)–(7) include cohorts 2011–2014.

Figure 1: Distribution of Parental Income by Free/Reduced Lunch Status



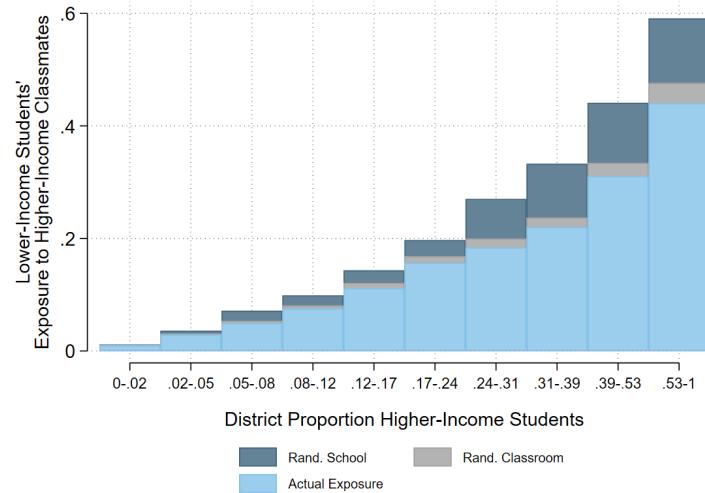
*Notes:* Parental Income is based on the adjusted parental gross contribution listed on financial aid applications from 2017 to 2022 linked to the cohort of students enrolled in grade 5 in 2012, 2013 and 2014. The vertical lines are the median income for each group. Outlier students with parents who have listed an income above 500,000 are top-coded so that we are better able to see the variation, but those observations are not top-coded when calculating the median and mean. Financial aid data are not available for 48%, 65% and 70% of students never, sometimes, and always on free/reduced lunch, respectively.

Figure 2: Exposure to Higher-Income Students: Observed Relative to Random Assignment



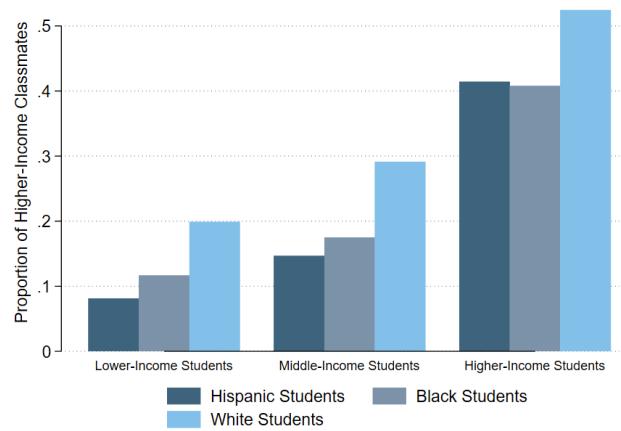
*Notes:* The bars capture the cumulative proportion of classmates that are higher-income calculated as the total number of higher-income students a given student is in a classroom with from grade 5 to expected grade 12 (excluding own status) divided by the total number of students in each classroom from grade 5 to expected grade 12 (excluding self). The average exposure is weighted by the number of students in each group. The dashed horizontal lines present the various benchmarks for integration. “S: Rand. Dist” lines capture the expected proportion of higher-income classmates had students been randomly assigned to districts in each year-cohort: district integration benchmark. The “S: Rand. Schl” lines capture the expected proportion of higher-income classmates had students been randomly assigned to schools within a district in each year-cohort (holding constant district composition): school integration benchmark. The “S: Rand. Class” lines capture the expected proportion of higher-income classmates had students been randomly assigned to classrooms within a school in each year-cohort (holding constant school composition): classroom integration benchmark.

Figure 3: The Percentage Point Gap in Exposure to Higher-Income Students Increases as the Proportion of Higher-Income students in the District Increases



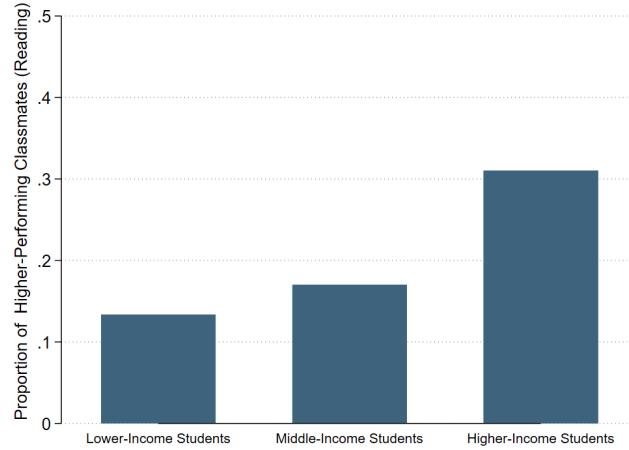
*Notes:* This plot presents lower-income students' exposure to higher-income students in districts with various proportions of higher-income students. Exposure is defined as the average proportion of higher-income classmates in a year. The light blue bar shows the observed proportion of higher-income students in lower-income students' classrooms. The navy bar presents the expected proportion of higher-income students had students been randomly assigned to schools in the district: school integration benchmark. The gray bar presents the expected proportion of higher-income students had students been randomly assigned to classrooms in the school: classroom integration benchmark. Districts are split into ten percentiles based on the distribution of the proportion of higher-income students in the district in a given school-year. The distribution is weighted by the number of students in each district (independent of income status). The lower and upper bound for each of the percentiles is shown on the x-axis.

Figure 4: Gap in Exposure to Higher-Income Students Exists Independent of Students' Race/Ethnicity



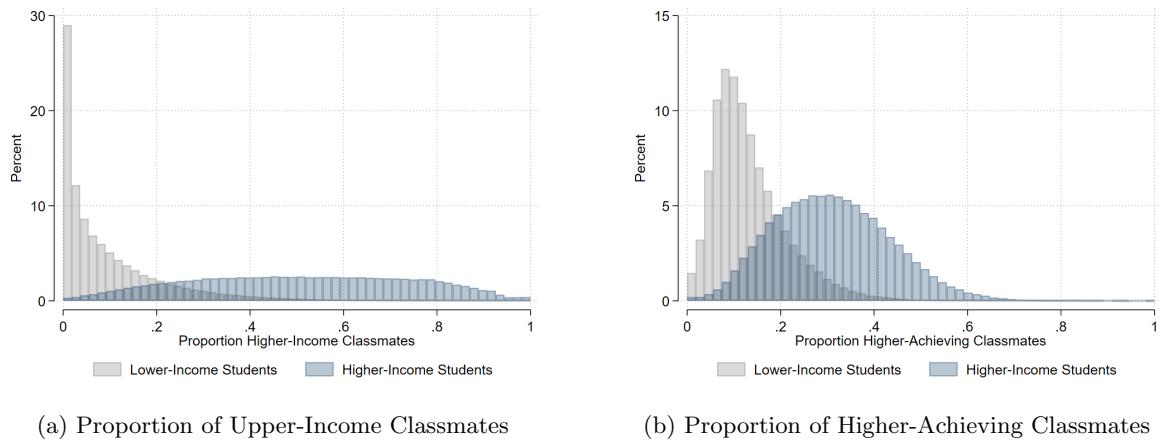
*Notes:* This plot captures the proportion of higher-income classmates across grades 5 to 12 for each income and racial/ethnic group. Higher-income students are defined as students never on free/reduced lunch.

Figure 5: The gap in Exposure to Higher-Performing (Top 24 Percentiles) Students is Smaller than the Gap in Exposure to Upper-Income Students



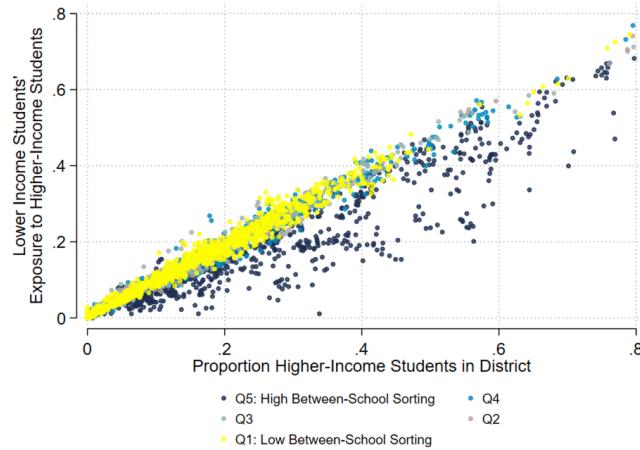
*Notes:* This plot captures students' cumulative proportion of higher-performing classmates between grade 5 to expected grade 12. I define higher-performing students as students who performed in the top 24 percentiles of test score based on the grade 4 standardized reading test. The percentiles are based on the distribution of students who have a grade 4 test score. Approximately 5% of students are missing grade 4 reading test scores.

Figure 6: At the 25th Percentile, Lower-Income Students Have 1% Upper-Income and 7% Higher-Achieving Classmates



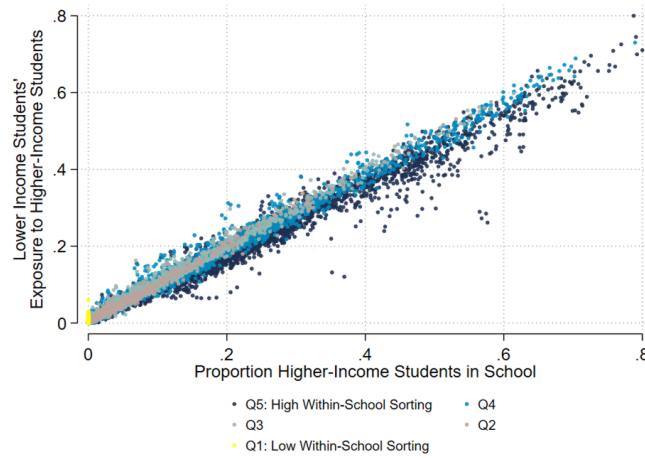
*Notes:* Every time a student is in a classroom with an upper-income/higher-achieving student, this is counted as one interaction. Histograms present the distribution of the proportion of classmates in cohorts 2012–2014 who are upper-income/higher-achieving between G5-12 for higher- and lower-income students. The average number of classroom interactions between grades five and twelve is approximately 2000.

Figure 7: Conditional on District Composition, Lower-Income Students' Exposure to Higher-Income Classmates



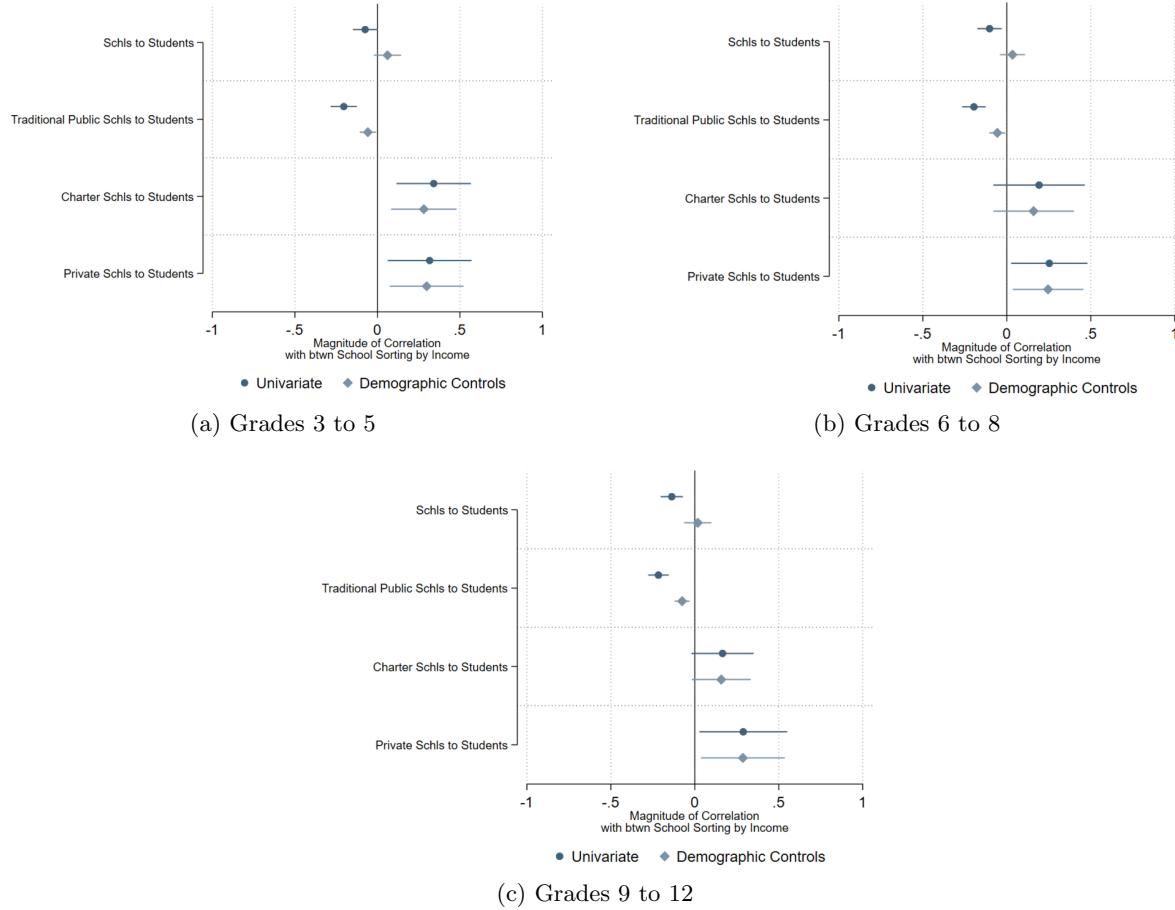
*Notes:* Each dot represents a district-grade. I only include districts with more than 50 lower-income students enrolled in a grade across three cohorts. The darker color represents district-grades with more sorting between schools, while the lighter dots represent district-grades with less sorting between schools. Sorting is captured by the variance ratio.

Figure 8: Conditional on School Composition, Lower-Income Students' Exposure to Higher-Income Classmates



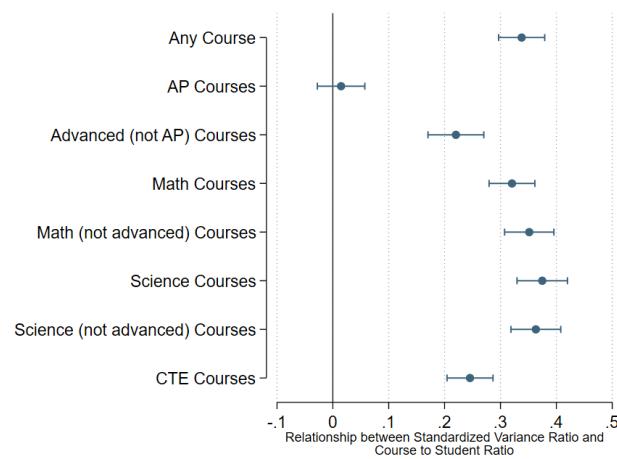
*Notes:* Each dot represents a school-grade. I only include schools with more than 50 lower-income students enrolled in a grade across three cohorts. The darker color represents school-grades with more sorting within school, while the lighter dots represent school-grades with less sorting within schools. Sorting is captured by the variance ratio.

Figure 9: Between-School Sorting in Public Schools is Higher in Districts with More Charter and Private School Options



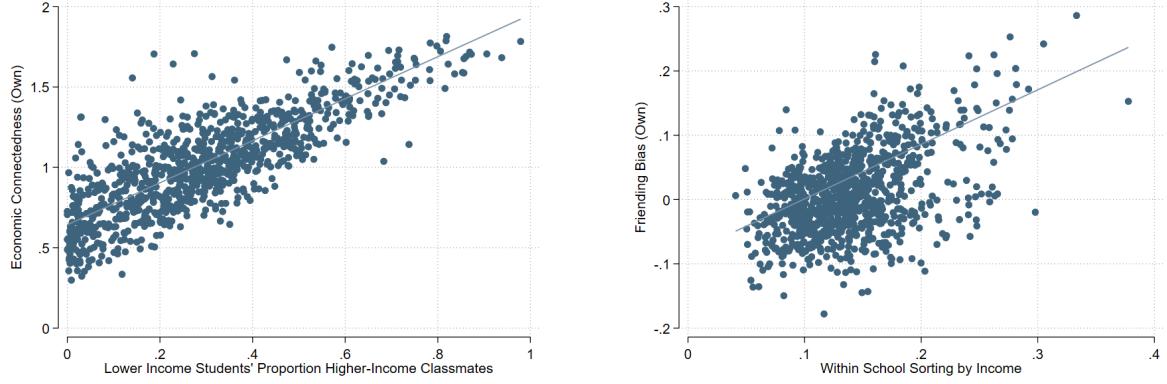
*Notes:* Both the number of schools to students and the variance ratio are standardized based on the school-to-student distribution and variance ratio in a district, weighted by the number of students enrolled and calculated within the grade group. Schools are placed into districts, and the ratio is calculated as the number of schools in the district to the number of students served in the district in 2019. The number of private schools, number of students and number of grades served in each private school are based on the Texas Alliances Accredited Private Commission archive data for the 2018–2019 school year. The “univariate” estimate is based on regressing the variance ratio on the standardized number of schools to students. The “demographic controls” estimate is based on regressing the variance ratio on the standardized number of schools to students, with controls for the proportion of Black, Hispanic and higher-income students in the district.

Figure 10: Within-School Sorting Correlates More with the Number of Science and Math Courses than Advanced Courses



*Notes:* Based on 2019 high school classroom enrollment data. School-level univariate regression coefficients from regressing the standardized sorting measure with the number of courses offered to students served are standardized in the first row. The other rows capture the correlation with other specific groups of courses to students served. Math courses capture the number of any math course to the number of students served, standardized. Math (not advanced) courses capture all math courses that are not advanced to the number of students served. The standardization is based on the weighted distribution of courses to students and within-school sorting across schools. The correlation is weighted by the number of students enrolled in a school. The 95% confidence intervals are presented. The standard errors are clustered at the school level.

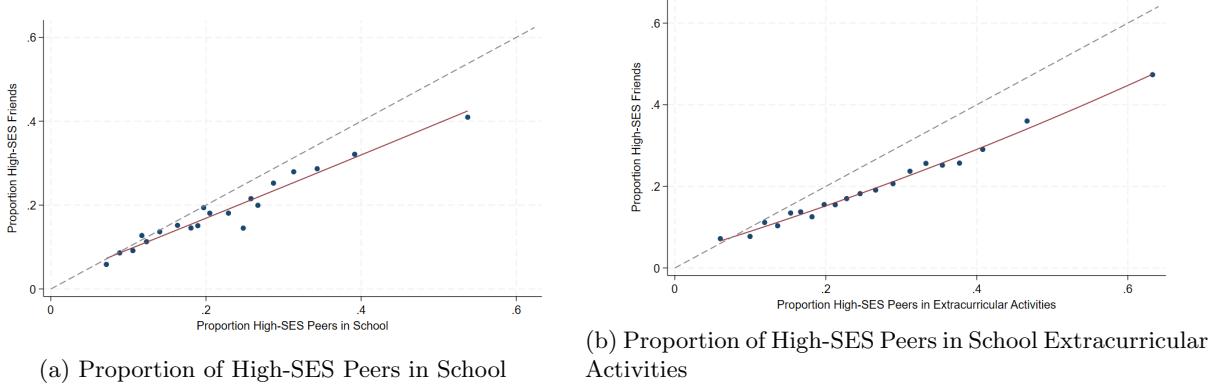
Figure 11: Strong Relationship between Likelihood of Befriending Higher-Income Students and Classroom Exposure



(a) Classroom Exposure and Economic Connectedness      (b) Classroom Sorting by Income and Friending Bias

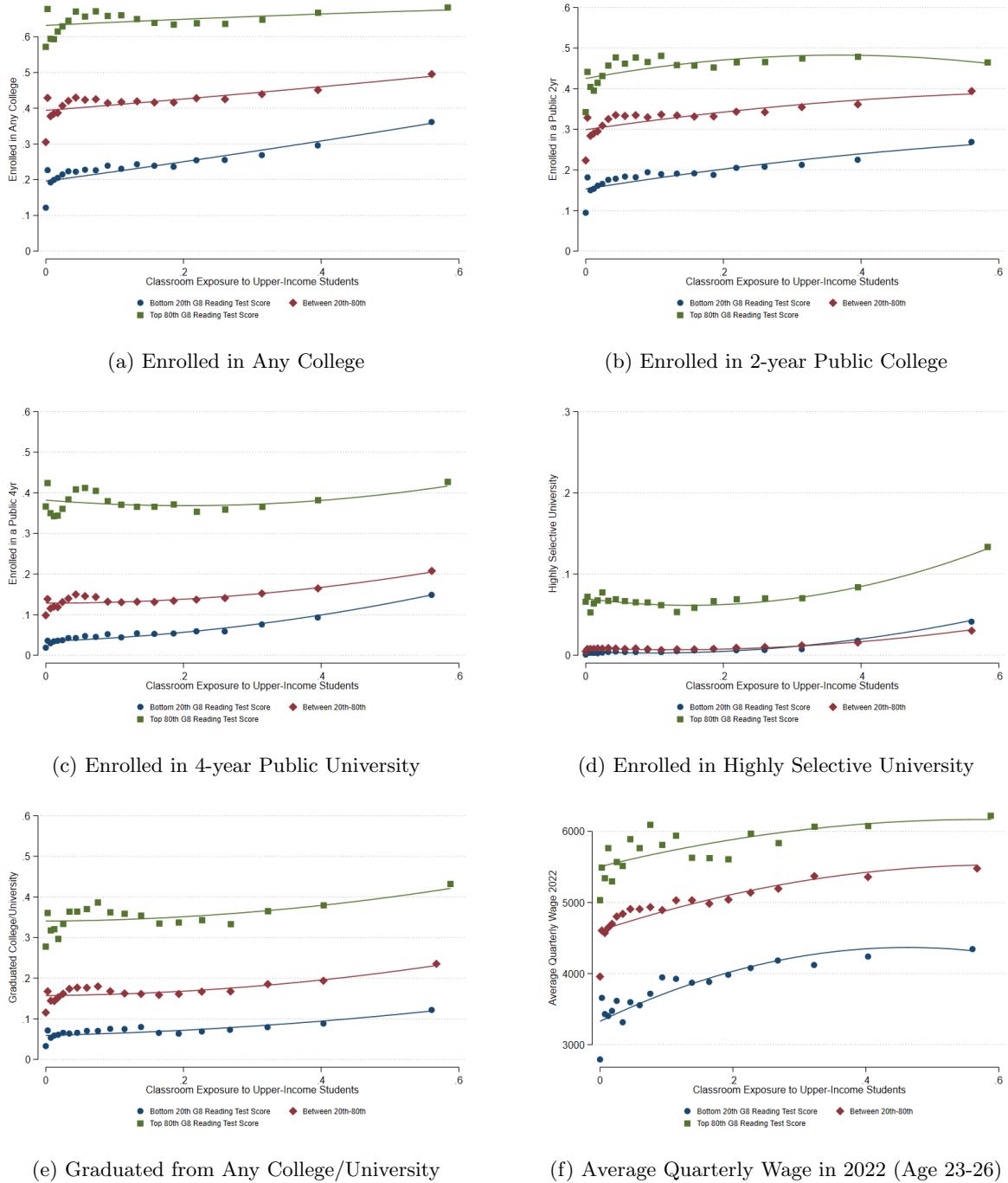
*Notes:* Each dot represents a school. I was able to link 1132 high schools to the friendship formation high school measures in Chetty et al. (2022). Of the 1132 schools, 981 have data on economic connectedness. I merged the two datasets with school name and district information. The fitted line weights schools by number of students enrolled. Panel (a) presents the correlation between high school measures of economic connectedness as captured by Chetty et al. (2022) and the proportion of higher-income students to whom lower-income students are exposed in each high school in 2012. Economic connectedness captures the likelihood of befriending a higher-income individual on Facebook. Panel (b) presents the correlation between high school measures of friending bias as captured by Chetty et al. (2022) and the difference in the proportion of higher-income students in higher- and lower-income students' classrooms, within a school (variance ratio). The variance ratio is based on 2012 student enrollment data. Friending bias captures the likelihood of befriending a higher-income individual on Facebook, conditional on the school proportion of higher-income students.

Figure 12: Strong Relationship Between Students' School and Extracurricular Composition and Friendship Pattern



*Notes:* The definition of high-SES is described in Section 3.2. The binned scatter plots are based on the full sample of low-SES students with data on friendship patterns. The bins are based on grouping the x-values into 20 equal-sized bins. Then, it computes the average y-variable value for each bin. The fitted lines are weighted by survey sampling weights. Panel (a) captures the relationship between a student's proportion of high-SES peers in school and their proportion of high-SES friends. Panel (b) captures the relationship between a student's proportion of high-SES peers in their school-listed extracurricular activities and their proportion of high-SES friends. In both panels, the dashed line is a 45-degree line representing a 1-to-1 relationship between the average proportion of high-SES students in school/extracurriculars and the average proportion of high-SES friends.

Figure 13: Exposure to Higher-Income Students, College Enrollment and Employment



*Notes:* The fitted quadratic line is based on the relationship between students' classroom exposure to upper-income classmates on the various outcomes, weighted by the number of students enrolled in each subgroup. College enrollment outcomes include cohorts 2011 to 2016. College graduation and wage outcomes include cohorts 2011 to 2014. The bins are based on grouping the x-values into 20 equal-sized bins. Then, it computes the average y-variable value for each bin.

Figure 14: Exposure to Higher-Income Students and School Resources

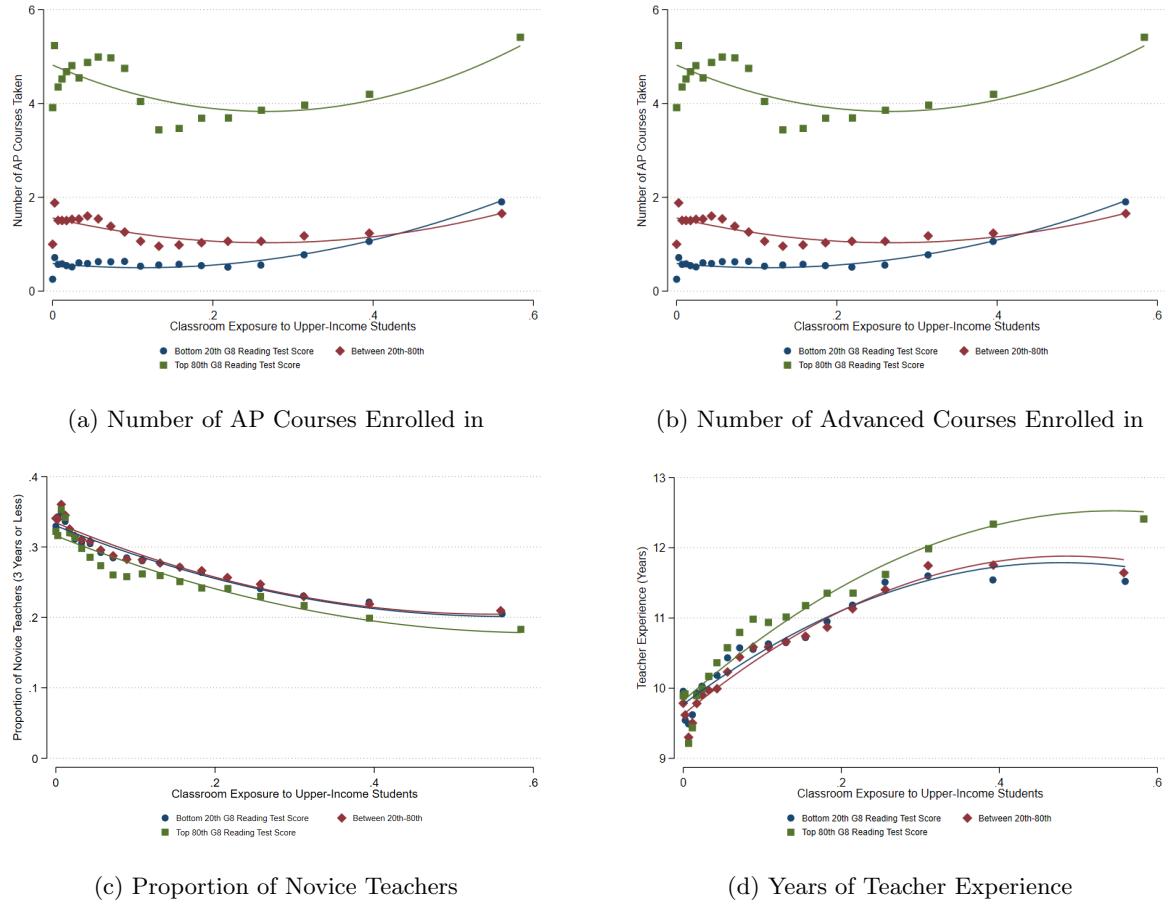
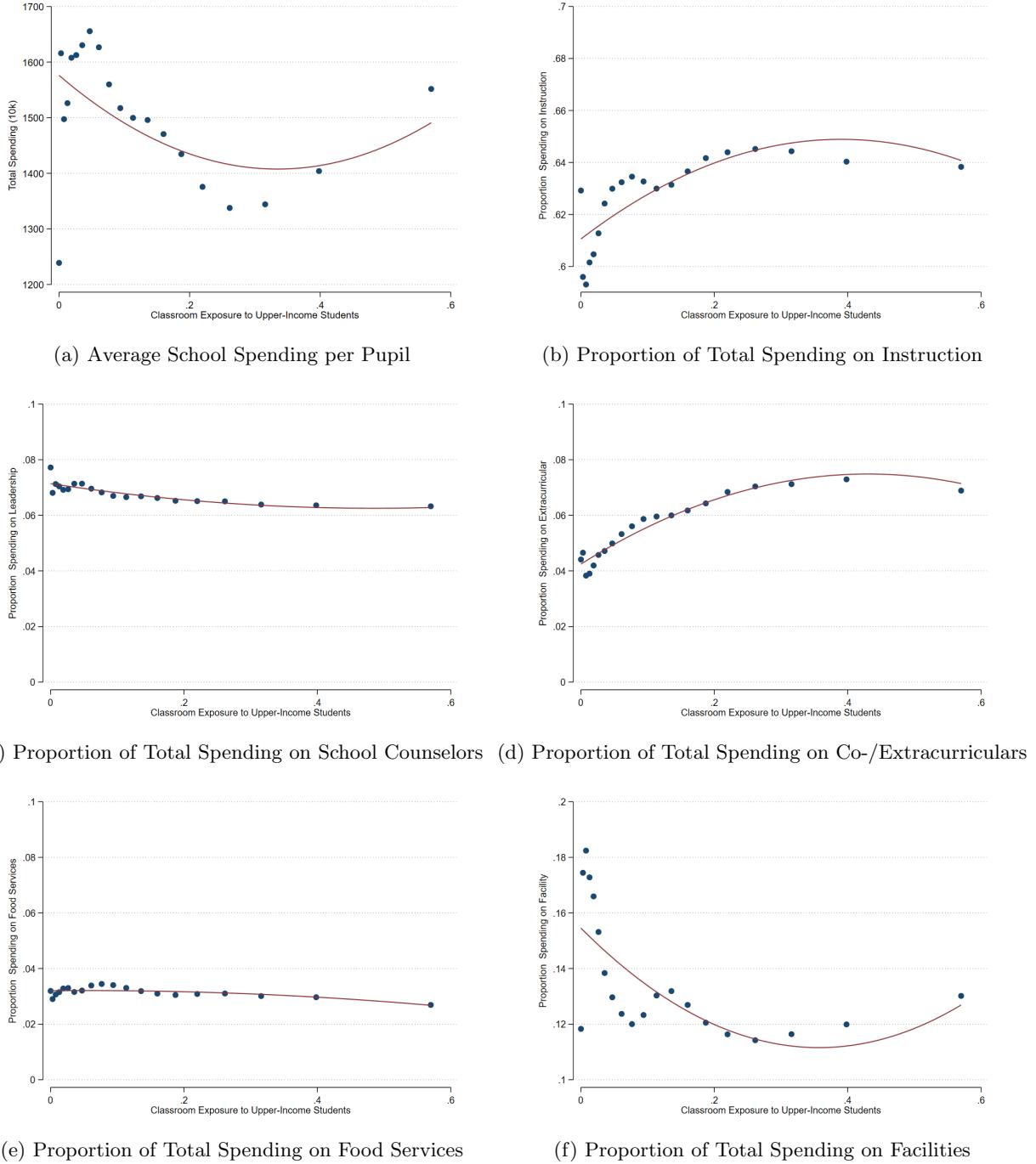
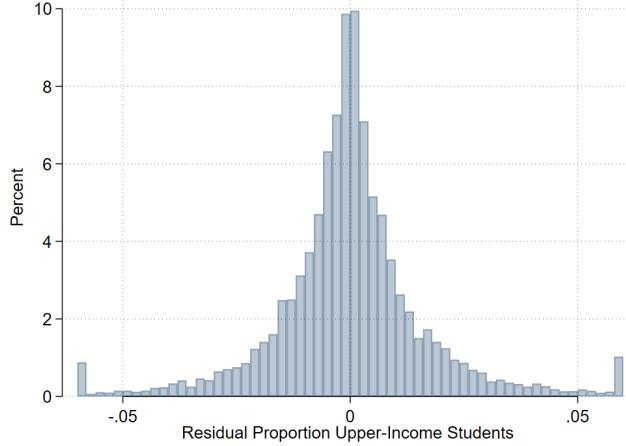


Figure 15: Exposure to Higher-Income Students and School Spending per Pupil



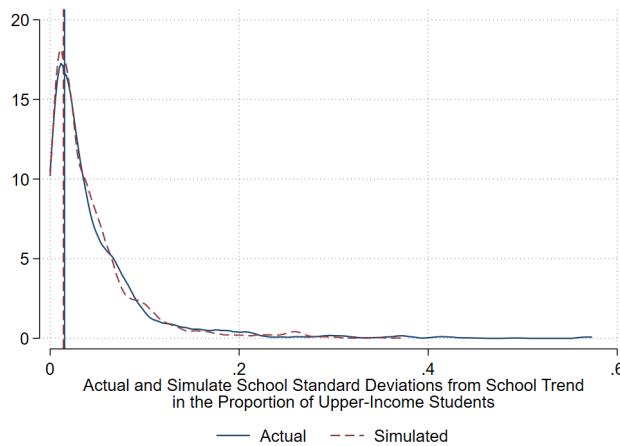
*Notes:* Spending is based on school actual spending data between 2012 and 2019 in each year divided by the total number of students enrolled in the school that year. Panel (a) captures the average yearly spending per pupil during the cohort's high-school years. Panels (b)–(f) capture the proportion of total spending on each category. The instruction category combines the following functions: instruction, instruction resources and media services, curriculum and staff development, and instructional leadership. Facilities combines the following functions: facility acquisition and construction and facility maintenance/operations. With the exception of food services, I only included categories that make up more than 4% of the budget, on average. The data include cohorts 2012 to 2016. The bins are based on grouping the x-values into 20 equal-sized bins. Then, it computes the average y-variable value for each bin.

Figure 16: Distribution of Residual Variation in Proportion Upper-Income Students



*Notes:* The histogram presents the variation in the proportion of upper-income students based on the residual from a regression with school and cohort fixed effects, as well as school time trends.

Figure 17: Simulated and Actual Distribution of within School Standard Deviations from School Trend



*Notes:* The figure captures the distribution of each school's standard deviation in the proportion of upper-income students from school trend (solid line) and each school's simulated average standard deviations in the proportion of upper-income students from school mean (dashed line). The vertical lines capture each distributions' median standard deviation in the proportion of upper-income students.

## 10 Supplementary Appendix

### 10.1 Appendix A: Notes on Simulation Calculations

To calculate the expected exposure to higher-income students for each income group under various random assignment levels, excluding own status, I do the following. I calculate the expected number of (higher-income) peers in each classroom had students been randomly assigned to classrooms in unit  $j$ . To calculate the expected number of students in each classroom, I divide the total number of student-classroom enrollments in Texas state public schools in unit  $j$  in each year-cohort  $TotalEnrollments_j$  by the total number of classrooms offered in unit  $j$  ( $N_j$ ). To calculate the expected number of higher-income students, I calculate the total number of student-classroom enrollments for higher-income students in unit  $j$  in each year-cohort  $TotalHighIncomeEnrollments_j$  and divide it by the total number of classrooms offered in unit  $j$  ( $H_j$ ). The choice of unit  $j$  depends on the random assignment level (state/district/school). This is shown in Equation (3)).

$$\begin{aligned} H_j &= \frac{TotalHighIncomeEnrollments_j}{TotalClassrooms_j} \\ N_j &= \frac{TotalEnrollments_j}{TotalClassrooms_j} \end{aligned} \tag{3}$$

For each student, her expected proportion of higher-income peers had students been randomly assigned to classrooms in unit  $j$  is equal to the proportion of higher-income students in unit  $j$ , i.e.,  $\frac{H_j}{N_j}$  adjusted to exclude the student herself. To adjust the proportion to exclude a student's own status, if a student is a higher-income student, I subtract from both the denominator and numerator her number of classroom enrollments  $n_{ij}$  during a given year in unit  $j$ . If a student is lower-income, I subtract from the denominator only her number of classroom enrollments as shown in Equations (4)) and (5)).

$$ExposureRand_{Hj} = \frac{H_j - n_{ij}}{N_j - n_{ij}} \tag{4}$$

$$ExposureRand_{Lj} = \frac{H_j}{N_j - n_{ij}} \tag{5}$$

To calculate the cumulative expected exposure for a given student across expected grades, I calculate the average of her weighted proportion of higher-income classmates across the years. The weights are based on the number of classrooms a student is enrolled in each year.

## 10.2 Appendix B: Variance Ratio

To capture within-unit (district or school) sorting by income, conditional on the proportion of upper-income students, I use the variance ratio (also known as the captures the correlation ratio and eta squared). The variance ratio the difference (gap) in exposure to higher-income students between higher- and lower-income students. If students are randomly assigned to districts, schools, and classrooms, we should expect higher- and lower-income students to have the same average number of high-income students. The variance ratio builds on the exposure measure, with a system-wide composition adjustment, as shown in Equation (6). To maintain the variance ratio interpretation of the measure being equal to 0 if students are enrolled in the same classroom, in this calculation (as done in prior literature), I include students' own income status.

$$\begin{aligned} VarianceRatio_j &= \frac{E[Prophighincome_{ij}|H = 1] - Prophighincome_j}{1 - Prophighincome_j} \\ &= E[Prophighincome_{ij}|H = 1] - E[Prophighincome_{ij}|H = 0] \end{aligned} \tag{6}$$

The variance ratio has a straightforward interpretation. In a perfectly integrated unit, the difference in exposure to higher-income students by income status would be equal to 0. In other words, everyone would be exposed to the same proportion of higher-income students in the school  $Prophighincome_j$ . In a perfectly segregated school, higher-income students would only be exposed to themselves  $E[Prophighincome_{ij}|H = 1] = 1$ , so the variance ratio would be equal to 1. Monarrez, Kisida and Chingos (2022) use the variance ratio to capture the impact of the expansion of charter schools on school segregation by race.

The variance ratio is sensitive to the unit's proportion of higher-income students: the size of the potential gap between groups differs by the units' composition (James and Taeuber, 1985). It is also sensitive to the number of classrooms/schools  $i$  within schools/districts  $j$  (Monarrez, Kisida and Chingos, 2022). A unit with more areas relative to the number of students has more scope for income sorting.

### 10.3 Appendix C: Segregation by Income and District School Offerings

To capture the relationship between school sorting and the number and type of schools offered, for each district, I calculate the number of schools that offer a given grade level and divide it by the number of students enrolled in that grade and district in 2019. Then, I run a regression with the number of schools to students on the level of between-school sorting as the outcome of interest. To capture school heterogeneity, I split schools into three types: traditional public schools, charter schools, and private schools.<sup>62</sup> I only observe students enrolled in public schools (traditional and charter), and so in the case of private schools, I examine the relationship between the number of private schools and sorting between public schools. Private schools provide higher-income students with an outside option. In doing so, they can alter both the peer composition of public schools and the policies and offerings of public schools. For simplicity, I focus on students enrolled in public schools (both traditional and charter) in the 2019 academic year in Texas.

Estimates of the relationship between the ratio of schools offered and the level of sorting by income is shown in Figure 9, with and without demographic controls. The errors are clustered at the district level and the estimates are weighted by the number of students enrolled.

The number of school options can also impact the level of within-school sorting by increasing competition for upper-income (or higher-achieving) students. Theoretically, public schools can use tracking to retain higher-performing students (Epple, Newlon, and Romano, 2002). Under tracking, higher-income students may be less likely to leave public schools to attend private schools. Figlio and Page (2002) find that higher-performing students are more likely to select into tracked schools. Similarly, Domina et al. (2017) find that schools with more advantaged students are more likely to increase tracking in response to policy pressures. As such, I run the same regression with within-school sorting as the outcome instead of between-school sorting. I find no evidence that within-school sorting in public schools is higher when the ratio of private schools to students is higher. There seems to be some evidence of a slightly lower level of within-school sorting in districts with more charter schools. These patterns are shown in Figure A29. Charter schools generally seem to have lower levels of within-school sorting, as shown in Table A4.

---

<sup>62</sup>The number of private schools in the district is based on the list of schools in 2018–2019 published by the Texas Private School Accreditation Commission. This list includes the number of students enrolled in each school and the grades served by the school. However, I do not know the number of students served in each grade or their demographics. I estimate the number of students in each grade by assuming that the same number of students are served in each grade. The within- and between-school sorting measures are based on students enrolled in public schools only.

## 10.4 Appendix D: Add Health Data and Income Measure

The at-home survey has information on parental income but only for a random subsample of students from the wider in-school survey. Therefore, I use the socioeconomic variables available for all students to build a model that predicts student income to capture all student SESs who share the same school extracurricular. I use a simple linear regression and use the following variables to predict student income: Father and Mother's occupation, Father and Mother's education, Father and Mother's employment status, number of individuals in the household, and missing indicators. I tested and built this model within the subset of students who were surveyed at home for whom I have income data. Then, I used this model to predict the income for the entire sample of students with the in-school survey. The students were then divided into three income groups based on their predicted income that mimic the income group sizes in the Texas data. High-SES students are those whose predicted income is in the top 24 percentiles, and low-SES students are students whose predicted income is in the bottom 29 percentiles. These three income indicators seem to effectively capture the variation in parental income. The median income is 26.6, 40.4 and 67.7 thousand dollars for low-, middle-, and high-SES students, respectively.<sup>63</sup> I summarize the characteristics of each of the income groups (including parental education) in Table 10.5.

The Add Health data list 33 extracurricular activities that students can check. The extracurricular activities listed are the following: yearbook, student council, honor society, newspaper, wrestling, volleyball, track, tennis, swimming, soccer, ice hockey, football, field hockey, basketball, baseball/softball, other sport, orchestra, chorus or choir, cheerleading/dance team, band, science club, math club, history club, Future Farmers of America, drama club, debate team, computer club, book club, Spanish club, Latin club, German club, French club, other club or organization.

---

<sup>63</sup>The median is based on the full sample of students in the at-home survey for whom I have parental income data. I list these in 1994 dollars, which are equivalent to 46.45, 70.5 and 118.21 thousand 2020 dollars. I split the set of students surveyed at home into two equally sized random groups. The in-sample half is used to build the model, and the out-of-sample half is used to test the model. The average income for out-of-sample high-SES students surveyed at home is 68.6 thousand dollars.

## 10.5 Appendix E: Tables and Figures

Table A1: Add Health Student Summary Statistics

Variable	Mean
Hispanic	0.1190
White	0.5952
Black	0.2034
Either Parent with College Degree	0.3632
Either Parent in Prof. Occupation	0.4444
Parental Income	43.1524
Median Income	35.0000
Proportion High-SES	0.2416
Proportion Low-SES	0.2872
Number of Students	85627
Number of Schools	142

*Notes.* This table summarizes demographic statistics for students surveyed in the Add Health in-school surveys. The variables “Median Income” and “Parental Income” are based on the at-home surveys of a random subset of students (17,238) for whom there are income data. The high- and low-SES indicators are based on the predicted income measure described in Section 3.2. All averages are weighted by survey sampling weights to be nationally representative.

Table A2: Add Health Student Summary Statistics by Income Group

Variable	Low-SES	Mid-SES	High-SES
Hispanic	0.1686	0.1175	0.0630
White	0.4105	0.6354	0.7362
Black	0.3183	0.1763	0.1198
Either Parent with College	0.0601	0.2718	0.8014
Either Parent in Prof. Occupation	0.1332	0.3569	0.8443
Proportion High-SES Friends	0.1715	0.2371	0.4034
Proportion High-SES in Extracurricular	0.2349	0.2734	0.3888
Proportion High-SES in School	0.1956	0.2276	0.3237
Number of Extracurriculars	1.9098	2.1660	2.8513
Deviation of Extracurricular from School Prop High-SES	0.0363	0.0435	0.0586
Parental Income (Thousands of Dollars)	26.6107	40.4065	67.7105
Median Income (Thousands of Dollars)	20.0000	36.0000	55.0000
Number of Students	23800	38961	22866

*Notes.* This table summarizes demographic statistics for students surveyed in the Add Health in-school surveys. The variables “Median Income” and “Parental Income” are based on the at-home surveys of a random subset of students (17,238) for whom there are income data. The high- and low-SES indicators are based on the predicted income measure described in Section 3.2. All averages are weighted by survey sampling weights to be nationally representative.

Table A3: Student School and Classroom Enrollment by Expected Grade

Variable	5	6	7	8	9	10	11	12
Enrolled in Charter School	.037 (.188)	.057 (.232)	.058 (.235)	.06 (.238)	.054 (.226)	.058 (.235)	.057 (.233)	.053 (.224)
Enrolled in Alternative School	.006 (.08)	.009 (.094)	.011 (.105)	.015 (.122)	.024 (.153)	.036 (.187)	.045 (.207)	.042 (.2)
Enrolled in Title 1 School	.706 (.455)	.612 (.487)	.568 (.495)	.57 (.495)	.418 (.493)	.413 (.492)	.421 (.494)	.435 (.496)
Special Education	.089 (.285)	.091 (.288)	.09 (.286)	.089 (.284)	.087 (.281)	.085 (.279)	.083 (.276)	.083 (.275)
Gifted Program	.094 (.292)	.097 (.296)	.099 (.298)	.099 (.299)	.096 (.294)	.095 (.294)	.098 (.298)	.105 (.306)
Enrolled in Bilingual Program	.171 (.377)	.08 (.271)	.069 (.254)	.071 (.256)	.075 (.263)	.078 (.268)	.063 (.244)	.035 (.185)
Number of Advanced Courses Taken	0 (0)	0 (.006)	.001 (.041)	.026 (.237)	.332 (.804)	.69 (1.345)	1.721 (2.709)	1.675 (2.264)
Number of Classrooms	9.889 (3.968)	9.396 (3.764)	9.70 (3.688)	10.077 (3.5)	14.767 (3.148)	14.861 (3.214)	12.095 (4.595)	8.481 (4.308)
Number of Courses Taken	9.277 (3.615)	8.632 (3.32)	8.806 (3.248)	9.288 (3.112)	14.167 (2.823)	14.333 (2.868)	11.695 (4.179)	8.232 (3.906)
Any Vocational Ed	.055 (.229)	.081 (.273)	.207 (.405)	.377 (.485)	.691 (.462)	.797 (.402)	.832 (.374)	.842 (.365)
Number of Students	1124550 1093343	1076510 1064386	1052391 1035723	988132 905982				

*Notes.* This table describes average enrollment across three cohorts of students followed from grade 5 to 12 from 2012 to 2021 in each type of school each year as well as a description of the course offerings in each type of school. The number in brackets is the standard deviation from the mean.

Table A4: School Classroom and Course Offerings by Grade

Variable	5	6	7	8	9	10	11	12
Number of Classrooms	63,262 (46,286)	134,705 (83,599)	178,256 (106,223)	197,025 (119,956)	798,786 (458,365)	1001,428 (528,681)	909,196 (583,499)	615,465 (454,558)
Number of Courses	10,387 (4,177)	17,052 (7,503)	24,806 (11,134)	31,569 (12,549)	137,994 (55,622)	208,034 (78,966)	232,223 (112,296)	193,952 (106,083)
Number of Advanced Courses	0	.002	.018	.286	6,527	19,015	34,147	37,771
Number of CTE Courses	(0)	(.065)	(.18)	(.749)	(5,351)	(12,55)	(22,686)	(24,838)
Number of Students	10,386 (4,176)	16,876 (7,374)	23,395 (10,321)	28,614 (11,274)	105,311 (41,72)	148,262 (56,57)	157,305 (76,935)	126,841 (70,561)
N. of Traditional Public Schools in District (weighted)	32,77 (42,932)	12,609 (14,626)	11,06 (13,307)	11,259 (13,371)	9,371 (11,566)	8,925 (11,043)	8,421 (10,514)	7,896 (10,169)
N. of Traditional Public Schools in District (unweighted)	3,727 (10,706)	2,128 (4,424)	1,926 (3,549)	1,965 (3,643)	1,857 (3,007)	1,851 (2,954)	1,79 (2,821)	1,687 (2,654)
Number of Schools	4263	2719	2423	2455	2309	2269	2206	2102
Number of Districts	1027	1023	1014	1013	981	979	979	979

*Notes.* This table summarizes the number and types of courses provided on average in schools serving students in the three cohorts in each grade level. Averages are weighted by the number of students served in each school. Numbers in brackets present the standard deviations from the mean. The average across cohorts in each grade level is based on the year a cohort is expected to be in that grade level

Table A5: Gap in Exposure by District Type [NCES]

Type of District	(1) Prop LI	(2) Prop HI	(3) Exp LI to HI	(4) VR: btwn Schl	(5) VR: within Schl	(6) Exp Gap [in Dist]	(7) Exp Gap [in Schl]
City-Large	.364	.149	.056	.154	.07	.042	.005
City-Midsize	.295	.204	.09	.094	.086	.037	.009
City-Small	.348	.226	.108	.099	.101	.041	.011
Rural-Distant	.208	.258	.191	.003	.083	.009	.009
Rural-Fringe	.231	.287	.146	.021	.082	.016	.009
Rural-Remote	.255	.187	.149	.002	.083	.005	.005
Suburb-Large	.225	.304	.114	.102	.08	.047	.008
Suburb-Midsize	.243	.273	.144	.025	.087	.017	.01
Suburb-Small	.197	.263	.195	.073	.112	.063	.024
Town-Distant	.293	.157	.128	.005	.096	.011	.01
Town-Fringe	.259	.188	.132	.011	.085	.011	.01
Town-Remote	.331	.135	.103	.006	.098	.009	.009

*Notes.* Each column presents the variable average weighted by the number of students enrolled in each district type. District type is based on NCES categorization. Higher- and lower-income students are defined as students never and always on free/reduced lunch, respectively. Column (3) captures the average cumulative proportion of higher-income students in lower-income students' classrooms. Column (4) captures the variance ratio between schools in a district, which is the level of between-school sorting. Column (5) captures the variance ratio within a school, which is the level of within-school sorting. The variance ratio is the difference in the proportion of higher-income students in higher- relative to other-income students' schools (classrooms) in a district (school). Column (6) captures the difference between average exposure to higher-income students shown in Column (3) and expected exposure had students been randomized to schools and classrooms in the district. Column (7) captures the difference between average exposure to higher-income students shown in Column (3) and expected exposure had students been randomized to classrooms in the school. The numbers in brackets are the standard deviation from the mean.

Table A6: Gap in Exposure by District Type [TEA]

Type of District	(1) Prop LI	(2) Prop HI	(3) Exp LI to HI	(4) VR: btwn Schl	(5) VR: within Schl	(6) Exp Gap [in Dist]	(7) Exp Gap [in Schl]
Charter School Districts	.329	.116	.047	.052	.05	.017	.002
Independent Town	.315	.161	.124	.01	.109	.016	.013
Major Suburban	.211	.297	.123	.114	.084	.054	.009
Major Urban	.384	.128	.055	.193	.068	.048	.005
Non-metro Fast Growing	.171	.305	.242	.007	.076	.016	.009
Non-metro Stable	.281	.193	.121	.003	.087	.007	.007
Other Central City	.329	.231	.086	.097	.086	.038	.008
Other Central City Suburban	.282	.247	.116	.018	.082	.014	.009
Rural	.245	.201	.151	.003	.08	.005	.005

*Notes.* Each column presents the variable average weighted by the number of students enrolled in each district type. District type is based on the TEA categorization in the data. Higher- and lower-income students are defined as students never and always on free/reduced lunch, respectively. Column (3) captures the average cumulative proportion of higher-income students in lower-income students' classrooms. Column (4) captures the variance ratio between schools in a district, which is the level of between-school sorting. Column (5) captures the variance ratio within a school, which is the level of within-school sorting. The variance ratio is the difference in the proportion of higher-income students in higher- relative to other-income students' schools (classrooms) in a district (school). Column (6) captures the difference between average exposure to higher-income students shown in Column (3) and expected exposure had students been randomized to schools and classrooms in the district. Column (7) captures the difference between average exposure to higher-income students shown in Column (3) and expected exposure had students been randomized to classrooms in the school. The numbers in brackets are the standard deviation from the mean.

Table A7: Correlation between Proportion Upper-Income Students and Lower-Income students' Long-Term Outcomes for Lower-Income Students

	Any College	2yr Public	4yr Public	Selective	Grad Any	Wage 2022	Income Percentile
Proportion Upper-Income	0.0595 (0.0191)	0.0800 (0.0165)	0.0189 (0.0128)	0.0223 (0.00307)	0.0297 (0.0129)	1082.5 (133.5)	3.800 (0.671)
Observations	1013911	1013911	1013911	1013911	700248	700248	700248
N clusters	2147	2147	2147	2147	1995	1995	1995

Standard errors in parentheses. Standard error clustered at the school level. The coefficients are based on the raw (no controls), except for cohort fixed effects, relationship between each of the school resources and students' grade 9 cohort composition. College enrollment outcomes are based on cohorts 2010 to 2016. College graduation and wage outcomes are based on cohorts 2010 to 2014.

Table A8: Correlation between Proportion Upper-Income Students and Access to School Resources for Lower-Income students

	Class Sorting	Teacher Experience	Novice Teachers	Spending	Advanced Courses	AP Courses
Proportion Upper-Income	0.151 (0.00749)	3.777 (0.430)	-0.218 (0.0190)	-6789.6 (739.5)	-1.307 (0.289)	-0.891 (0.199)
Observations	841237	737009	737009	707296	1013911	1013911
N clusters	1987	1997	1997	1753	2147	2147

Standard errors in parentheses. Standard error clustered at the school level. The coefficients are based on the raw (no controls), except for cohort fixed effects, relationship between each of the school resources and students' grade 9 cohort composition. Class sorting is based on the variance ratio. Teacher experience is based on a student's average teacher experience between grades 9 and 12. A novice teacher is one with three or less number of years of experience. School spending is based on grades 9 to 12 average school spending per pupil. Advanced and AP courses are based on the number of each type of course a student took between grades 9 and 12. School spending and teacher experience outcomes are based on cohorts 2012 to 2016 based on data availability. School sorting based on cohorts 2011 to 2016 and advanced course taking outcomes based on 2010 to 2016 cohorts.

Table A9: Grade 3–8 Average Standardized Reading Scores for 9th Graders

Cohort	% Missing	All	High-Income	Low-Income	High-Achieve
2010	4.1	-.084	.361	-.396	.585
2011	4.4	-.059	.387	-.37	.608
2012	3.9	-.104	.385	-.421	.627
2013	4	-.069	.448	-.38	.746
2014	5.8	-.055	.401	-.325	.712
2015	5	-.065	.393	-.333	.729
2016	4.5	-.111	.408	-.396	.802

Table A10: Impact of an Increase in the Proportion of Higher-Income Students on Employment Industry for Lower-Income Students

	(1)	(2)	(3)	(4)	(5)
	Construction	Information	Finance	Prof. Science	Education
Proportion Upper-Income	0.00157 (0.0146)	-0.00707 (0.00577)	-0.0137 (0.00986)	0.00893 (0.0115)	0.00373 (0.0120)
N Clusters	1992	1992	1992	1992	1992
N	700245	700245	700245	700245	700245

Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (1). The models include cohorts 2010–2014.

Table A11: Impact of an Increase in the Proportion of Higher-Income Students on a Cohort’s Racial/Ethnic Composition

	(1)	(2)	(3)
	Prop. Hispanic	Prop. Black	Prop. White
Proportion Upper-Income	-0.243 (0.0155)	-0.0567 (0.00934)	0.322 (0.0150)
N Clusters	2141	2141	2141
N	1013905	1013905	1013905

Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (1). Models include cohorts 2010–2016.

Table A12: Impact of an Increase in the Proportion Upper-Income Peers on Lower-Income Students: Controlling for Higher-Achieving Students  
(Based on Math Test Score)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Any College	2yr Public	4yr Public	Selective	Grad Any	Wage 2022	Income Percentile
Proportion Upper-Income	0.0678 (0.0243)	0.0543 (0.0241)	0.0520 (0.0192)	0.00571 (0.00587)	0.0450 (0.0223)	1008.4 (395.9)	4.079 (1.653)
Proportion Higher-Achieving (Math)	-0.115 (0.0143)	-0.0864 (0.0128)	-0.0724 (0.0123)	-0.0137 (0.00349)	-0.0867 (0.0138)	-1122.6 (186.0)	-3.802 (0.827)
N Clusters	2141	2141	2141	2141	1992	1992	1992
N	1013905	1013905	1013905	1013905	700245	700245	700245

Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (2). Higher-Achieving Students are defined as those who scored in the top 25 percentiles based on the distribution of students' standardized math test scores. Models (1)–(4) include cohorts 2010–2016. Models (5)–(7) include cohorts 2010–2014.

Table A13: Impact of an Increase in the Proportion of Upper-Income Peers: Controlling for Higher-Achieving Students (Based on G8 Math and Reading Test Score)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Any College	2yr Public	4yr Public	Selective	Grad Any	Wage 2022	Income Percentile
Proportion Upper-Income	0.0787 (0.0240)	0.0600 (0.0242)	0.0613 (0.0188)	0.00767 (0.00580)	0.0462 (0.022)	1011.7 (394.4)	4.125 (1.657)
Proportion High-Achieve (Math+Reading)	-0.195 (0.0131)	-0.132 (0.0129)	-0.137 (0.0109)	-0.0271 (0.00363)	-0.0957 (0.0142)	-1178.9 (201.6)	-4.170 (0.849)
N Clusters	2141	2141	2141	2141	1992	1992	1992
N	1013905	1013905	1013905	1013905	700245	700245	700245

Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (2). Higher-achieving students are defined as those who scored in the top 25 percentiles based on the distribution of students' average standardized math and reading test scores. Models (1)–(4) include cohorts 2010–2016. Models (5)–(7) include cohorts 2010–2014.

Table A14: Impact of an Increase in the Proportion Upper-Income Peers on Lower-Income Students: Controlling for Higher-Achieving Students  
(Based on G3-8 Average Reading Test Score)

	(1) Any College	(2) 2yr Public	(3) 4yr Public	(4) Selective	(5) Grad Any	(6) Wage 2022	(7) Income Percentile
Proportion Upper-Income	0.0677 (0.0250)	0.0532 (0.0257)	0.0528 (0.0195)	0.00685 (0.00595)	0.0321 (0.0228)	894.1 (396.7)	3.815 (1.675)
Proportion High-Achieve (Reading G3-8)	-0.0899 (0.0208)	-0.0638 (0.0225)	-0.0601 (0.0160)	-0.0154 (0.00470)	-0.0195 (0.0196)	-460.3 (290.3)	-2.052 (1.245)
N Clusters	2141	2141	2141	2141	1992	1992	1992
N	1013905	1013905	1013905	1013905	700245	700245	700245

Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (2). Higher-Achieving Students are defined as those who scored in the top 25 percentiles based on the distribution of students' standardized reading test scores averaged between grades 3 and 8. Models (1)–(4) include cohorts 2010–2016. Models (5)–(7) include cohorts 2010–2014.

Table A15: Impact of Peer Income on Lower-Income Students, Conditional on Peer Test Scores (Continuous G8 Reading Test Score Measure)

	(1) Any College	(2) 2yr Public	(3) 4yr Public	(4) Selective	(5) Grad Any	(6) Wage 2022	(7) Income Percentile
Proportion Upper-Income	0.0564 (0.0246)	0.0454 (0.0247)	0.0453 (0.0194)	0.00432 (0.00585)	0.0320 (0.0226)	836.5 (399.2)	3.647 (1.657)
Peer Standardized G8 Test-Score	-0.0348 (0.00469)	-0.0256 (0.00533)	-0.0230 (0.00355)	-0.00394 (0.00104)	-0.00896 (0.00400)	-136.0 (6.33)	-0.790 (0.274)
N Clusters	2133	2133	2133	2133	1990	1990	1990
N	1013528	1013528	1013528	1013528	700060	700060	700060

Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (2). Average peer test score excludes own test score and is based on the grade 8 reading test score. Models (1)–(4) include cohorts 2010–2016. Models (5)–(7) include cohorts 2010–2014.

Table A16: Impact of an Increase in the Proportion of Upper-Income Students on Lower-Income Students, Conditional on Racial/Ethnic Composition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Any College	2yr Public	4yr Public	Selective	Grad Any	Wage 2022	Income Percentile
[1em] Proportion Upper-Income	0.0497 (0.0252)	0.0367 (0.0252)	0.0437 (0.0199)	0.00410 (0.00588)	0.0281 (0.0232)	838.8 (413.4)	3.204 (1.697)
Proportion Hispanic Students	0.0280 (0.0219)	0.0134 (0.0205)	0.0279 (0.0142)	0.00305 (0.00383)	0.00962 (0.0162)	264.9 (275.7)	0.0842 (1.219)
Proportion Black Students	-0.0449 (0.0334)	-0.0701 (0.0305)	-0.0174 (0.0238)	0.00548 (0.00638)	-0.0299 (0.0238)	-111.4 (432.4)	-2.269 (1.820)
N Clusters	2141	2141	2141	2141	1992	1992	1992
N	1013905	1013905	1013905	1013905	700245	700245	700245

Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (2) but interacted with the proportion of Hispanic and Black students instead. Models (1)–(4) include cohorts 2010–2016. Models (5)–(7) include cohorts 2010–2014.

Table A17: Impact of an Increase in the Proportion of Higher-Income Students on Lower-Income Students within a Racial/Ethnic Group: Hispanic Students

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Any College	0.0755 (0.0711)	0.0165 (0.0793)	0.0327 (0.0664)	0.0211 (0.0168)	0.0701 (0.0682)	1067.9 (952.5)	2.980 (4.065)
Proportion Upper-Income Hispanic Students	2013	2013	2013	2013	1885	1885	1885
N Clusters	688557	688557	688557	688557	475618	475618	475618
N							

Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (1). Models (1)–(4) include cohorts 2010–2016. Models (5)–(7) include cohorts 2010–2014. The sample is limited to Hispanic lower-income students.

Table A18: Impact of an Increase in the Proportion of Higher-Income Students on Lower-Income students within a Racial/Ethnic Group: White students

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Any College	0.0747 (0.0508)	0.0643 (0.0479)	0.0257 (0.0339)	0.0135 (0.0106)	-0.00389 (0.0445)	1962.6 (804.4)	13.12 (3.674)
Proportion Upper-Income White Students	1796	1796	1796	1796	1668	1668	1668
N Clusters	88135	88135	88135	88135	60843	60843	60843
N							

Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (1). Models (1)–(4) include cohorts 2010–2016. Models (5)–(7) include cohorts 2010–2014. The sample is limited to White lower-income students.

Table A19: Impact of an Increase in the Proportion of Higher-Income Students on Lower-Income Students within a Racial/Ethnic Group: Black Students

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Any College	2yr Public	4yr Public	Selective	Grad Any	Wage 2022	Income Percentile
Proportion Upper-Income Black Students	-0.185 (0.195)	-0.240 (0.178)	0.429 (0.160)	0.0216 (0.0466)	0.181 (0.179)	5631.2 (6198.7)	-4.042 (12.40)
N Clusters	1535	1535	1535	1535	1390	1390	1390
N	154185	154185	154185	154185	106733	106733	106733

Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (1). Models (1)–(4) include cohorts 2010–2016. Models (5)–(7) include cohorts 2010–2014. The sample is limited to Black lower-income students.

Table A20: Indirect Impact of the Proportion of Upper-Income Students: College Enrollment and Employment

	(1) Any College	(2) 2yr Public	(3) 4yr Public	(4) Selective	(5) Grad Any	(6) Wage 2022	(7) Income Percentile
Proportion Upper-Income (S)	0.0922 (0.0345)	0.0740 (0.0356)	0.0434 (0.0281)	0.00495 (0.00841)	0.0376 (0.0312)	398.9 (519.9)	1.578 (2.239)
Proportion Upper-Income District (-S)	0.00566 (0.0161)	0.0163 (0.0154)	-0.00396 (0.0103)	-0.00200 (0.00362)	-0.00903 (0.0179)	163.3 (246.2)	0.582 (1.085)
N Clusters	1559	1559	1559	1559	1369	1369	1369
N	878484	878484	878484	878484	604829	604829	604829

Standard errors are in parentheses and clustered at the school level, controlling for the proportion upper-income students in district schools other than the one a student is enrolled in. The proportion of high-achieving students are those who scored in the top 24 percentiles of the grade 8 reading test score. Models (1)–(4) include cohorts 2010–2016. Models (5)–(7) include cohorts 2010–2014.

Table A21: Indirect Impact of the Proportion of Upper-Income Students: Resources

	(1) Classroom Exposure	(2) Class Sorting	(3) Teacher Experience	(4) Novice Teachers	(5) Spending	(6) Advanced	(7) AP
Proportion Upper-Income (S)	0.520 (0.0132)	0.153 (0.0224)	0.215 (0.603)	0.0186 (0.0367)	-789.2 (807.2)	0.197 (0.503)	0.403 (0.345)
Proportion Upper-Income District (-S)	0.000650 (0.00344)	-0.00392 (0.00763)	-0.132 (0.233)	0.0177 (0.0146)	611.9 (665.1)	0.0503 (0.141)	0.0906 (0.0909)
N Clusters	1518	1383	1391	1391	1218	1559	1559
N	767610	733464	642580	642580	620458	878484	878484

Standard errors are in parentheses and clustered at the school level, controlling for the proportion upper-income students in district schools other than the one a student is enrolled in. Model (1) include cohorts 2011–2016 because the classroom data needed to calculate the within-school variance ratio are only available starting in 2011. Models (2)–(4) include cohorts 2012–2016 because teacher and school budget data linked to classrooms are only available starting in 2012. Models (5)–(6) include cohorts 2010–2016.

Table A22: Impact of Changes in Income and Achievement Composition on Lower-Income Students: 4 Groupings

	(1) Any College	(2) 2yr Public	(3) 4yr Public	(4) Selective	(5) Grad Any	(6) Wage 2022	(7) Income Percentile
Proportion Upper-Income and High-Achieve	-0.0331 (0.0337)	-0.00816 (0.0312)	0.00821 (0.0254)	-0.00956 (0.00862)	0.0315 (0.0309)	557.9 (520.6)	1.778 (2.129)
Proportion Upper-Income and Lower-Achieve	0.130 (0.0294)	0.0813 (0.0299)	0.100 (0.0218)	0.0125 (0.00687)	0.0288 (0.0270)	639.9 (452.9)	2.229 (1.958)
Proportion Low-Income and High-Achieve	0.0127 (0.0462)	-0.0393 (0.0403)	0.0539 (0.0428)	-0.00997 (0.0102)	-0.0545 (0.0427)	-1361.4 (503.9)	-6.289 (2.186)
Proportion Low-Income and Lower-Achieve	0.0463 (0.0193)	0.0335 (0.0203)	0.0553 (0.0141)	0.00323 (0.00408)	0.0169 (0.0179)	-207.7 (251.2)	-2.163 (1.130)
N Clusters	2141	2141	2141	2141	1992	1992	1992
N	1013905	1013905	1013905	1013905	700245	700245	700245

Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (1) but instead of the main independent variable (proportion upper-income students), students are split into four groups based on if they are upper-income and/or higher-achieving (test score in the top 24 percentiles of grade 8 reading test). Models (1)–(4) include cohorts 2010–2016 and Models (5)–(7) include cohorts 2010–2014.

Table A23: Impact of Increase in Proportion Upper-Income Students on Lower-Income Students: Excluding Outlier Observations (Limiting the Sample to School with Within School Standard Deviations less than 0.3pp)

	(1) Any College	(2) 2yr Public	(3) 4yr Public	(4) Selective	(5) Grad Any	(6) Wage 2022	(7) Income Percentile
Proportion Upper-Income	0.0451 (0.0247)	0.0370 (0.0246)	0.0378 (0.0194)	0.00308 (0.00583)	0.0272 (0.0228)	756.2 (400.8)	3.169 (1.657)
N clusters	2129	2129	2129	2129	1981	1981	1981
N	1013816	1013816	1013816	1013816	700175	700175	700175

Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (1) but without controlling for students grade 8 test and demographic characteristics. Models (1)–(4) include cohorts 2010–2016 and Models (5)–(7) include cohorts 2010–2014. Models exclude observations from schools with a standard deviation from school mean trend in the proportion of upper-income peers larger than 0.3pp.

Table A24: Proportion Upper-Income Peers and Lower-Income Students' Average Cohort Income

	Prop FAD	Parent Income (FAD)	Years FRPL	Prop FRPL
Proportion Upper-Income	0.247 (0.0193)	52.83 (4.669)	-6.761 (0.137)	-0.791 (0.0146)
Observations	1013905	1011522	1013905	1013905
Group Mean	0.400	40.02	6.350	6.350
N clusters	2141	2080	2141	2141

Standard errors clustered at the school level in parentheses. Each coefficient is based on a regression of the variable on the proportion upper-income students, including school and cohort fixed effects, and school time trends. The outcomes capture a lower-income students' average cohort characteristics excluding own status. "Prop FAD" captures the proportion of a student's cohort who submitted their financial aid (FAD) applications. "Parent Income (FAD)" is the average cohorts' income based on students' reported parental income in their FAD application. Years FRPL captures a student's peers average number of years on free/reduced lunch and prop. FRPL captures a student's peers average proportion of years on free/reduced lunch.

Table A25: Correlation Between Proportion Upper-Income Students in High-School and Lower-Income Students' Demographic Characteristics

	Proportion Upper-Income Students
Black Student	-0.0158 (0.0158)
Group Mean	0.15
N Clusters	2141
N	1013905
White Student	0.0250 (0.0171)
Group Mean	0.09
N Clusters	2141
N	1013905
Hispanic Student	-0.00684 (0.0225)
Group Mean	0.68
N Clusters	2141
N	1013905
Ever Immigrant	-0.01474 (0.0133)
Group Mean	0.10
N Clusters	2141
N	1013905
Female Student	-0.0101 (0.0246)
Group Mean	0.47
N Clusters	2141
N	1013905
Income Reported on FAD	-0.931 (4.434)
Group Mean	28.92
N Clusters	1926
N	298598

Standard errors clustered at the school level in parentheses. Each coefficient is based on a regression of the variable on the proportion upper-income students, including school and cohort fixed effects, and school time trends. Ever immigrant is based on if a student was ever assigned an immigrant status in the Texas data. A student is labeled as immigrant in the data if they were not born in any state and have not attended any state for more than three academic years.

Table A26: Impact of Proportion Upper-Income Students on Lower-Income Students' College Enrollment and wages: Robustness to Specification

	Model 1	Model 2	Model 3	Model 4	Model 5
<i>Enrolled in College/University</i>					
Proportion Upper-Income	0.116 (0.0241)	0.0633 (0.0250)	0.0665 (0.0247)	0.0454 (0.0246)	0.0538 (0.0301)
Group Mean	0.380	0.380	0.380	0.380	0.380
N Clusters	2141	2141	2141	2141	2141
<i>Enrolled in a Public 2yr College</i>					
Proportion Upper-Income	0.113 (0.0251)	0.0511 (0.0248)	0.0531 (0.0246)	0.0374 (0.0245)	0.0120 (0.0302)
Group Mean	0.290	0.290	0.290	0.290	0.290
N Clusters	2141	2141	2141	2141	2141
<i>Enrolled in a Public 4yr College</i>					
Proportion Upper-Income	0.0611 (0.0163)	0.0482 (0.0188)	0.0497 (0.0187)	0.0379 (0.0193)	0.0568 (0.0228)
Group Mean	0.140	0.140	0.140	0.140	0.140
N Clusters	2141	2141	2141	2141	2141
<i>Enrolled in a Selective University</i>					
Proportion Upper-Income	0.00902 (0.00467)	0.00481 (0.00583)	0.00474 (0.00582)	0.00304 (0.00581)	0.00737 (0.00728)
Group Mean	0.0200	0.0200	0.0200	0.0200	0.0200
N Clusters	2141	2141	2141	2141	2141
<i>Graduated College/University</i>					
Proportion Upper-Income	0.0614 (0.0202)	0.0397 (0.0223)	0.0400 (0.0222)	0.0272 (0.0227)	0.0238 (0.0274)
Group Mean	0.140	0.140	0.140	0.140	0.140
N Clusters	1992	1992	1992	1992	1991
<i>Average Quarterly Wage 2022</i>					
Proportion Upper-Income	812.3 (343.4)	987.1 (410.0)	984.3 (409.3)	777.3 (398.9)	796.3 (613.7)
Group Mean	4307.2	4307.2	4307.2	4307.2	4307.2
N Clusters	1992	1992	1992	1992	1991
<i>Income Percentile 2022</i>					
Proportion Upper-Income	2.501 (1.459)	4.298 (1.720)	4.308 (1.722)	3.297 (1.650)	2.926 (2.306)
Group Mean	47.69	47.69	47.69	47.69	47.69
N Clusters	1992	1992	1992	1992	1991
Linear School Time Trends		X	X	X	X
Student Demographic Controls			X	X	X
Student G8 Test Controls				X	X
Grade 8 School-Cohort Fixed Effects					X

Standard errors are in parentheses and clustered at the school level. Coefficients are based on Equation (1). All models include cohort and school fixed effects. College graduation and wages include cohorts 2010–2014. Group Mean is the average for lower-income students in the sample. Student demographic controls include student own race, immigrant status and gender. Student grade 8 test controls include her grade 8 math and reading scores and an indicator for missing grade 8 math and reading tests.

Table A27: Impact of Proportion Upper-Income and High-Achieving Students on Lower-Income Students' College Enrollment and wages: Robustness to Specification

	Model 1	Model 2	Model 3	Model 4	Model 5
<i>Enrolled in College/University</i>					
Proportion Upper-Income	0.0876 (0.0241)	0.0522 (0.0247)	0.0561 (0.0244)	0.0642 (0.0249)	0.0648 (0.0305)
Proportion High-Achieve	0.100 (0.0205)	0.0509 (0.0197)	0.0480 (0.0195)	-0.0854 (0.0208)	-0.0505 (0.0219)
Group Mean	0.380	0.380	0.380	0.380	0.380
N Clusters	2141	2141	2141	2141	2141
<i>Enrolled in a Public 2yr College</i>					
Proportion Upper-Income	0.0927 (0.0258)	0.0446 (0.0253)	0.0471 (0.0251)	0.0518 (0.0253)	0.0200 (0.0316)
Proportion High-Achieve	0.0716 (0.0202)	0.0298 (0.0201)	0.0276 (0.0198)	-0.0661 (0.0205)	-0.0368 (0.0235)
Group Mean	0.290	0.290	0.290	0.290	0.290
N Clusters	2141	2141	2141	2141	2141
<i>Enrolled in a Public 4yr College</i>					
Proportion Upper-Income	0.0395 (0.0159)	0.0404 (0.0187)	0.0420 (0.0186)	0.0481 (0.0194)	0.0737 (0.0228)
Proportion High-Achieve	0.0752 (0.0144)	0.0362 (0.0153)	0.0353 (0.0154)	-0.0461 (0.0159)	-0.0768 (0.0176)
Group Mean	0.140	0.140	0.140	0.140	0.140
N Clusters	2141	2141	2141	2141	2141
<i>Average Quarterly Wage 2022</i>					
Proportion Upper-Income	566.5 (343.1)	979.9 (412.8)	978.7 (412.2)	893.4 (401.6)	684.8 (624.0)
Proportion High-Achieve	851.8 (234.5)	35.72 (259.5)	32.68 (259.5)	-507.7 (252.3)	347.4 (335.6)
Group Mean	4307.2	4307.2	4307.2	4307.2	4307.2
N Clusters	1992	1992	1992	1992	1991
<i>Income Percentile 2022</i>					
Proportion Upper-Income	1.669 (1.468)	4.026 (1.729)	4.056 (1.731)	3.727 (1.665)	2.168 (2.337)
Proportion High-Achieve	2.885 (1.041)	1.236 (1.127)	1.162 (1.122)	-1.968 (1.081)	2.319 (1.502)
Group Mean	47.69	47.69	47.69	47.69	47.69
N Clusters	1992	1992	1992	1992	1991
Linear School Time Trends		X	X	X	X
Student Demographic Controls			X	X	X
Student G8 Test Controls				X	X
Grade 8 School-Cohort Fixed Effects					X

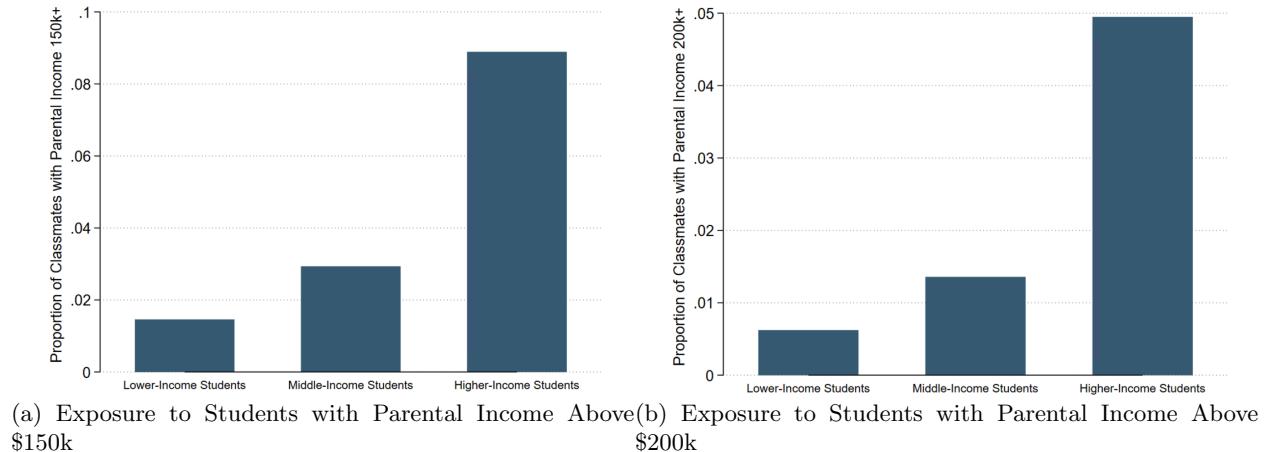
Similar to Table A26 but controlling for peer achievement as well based on average proportion high-achieving students based on distribution of grade 8 reading test-scores (excluding self)88

Table A28: Impact of Increase in Proportion Upper-Income Students: Robustness to Using G3 Reading Instead of G8 Reading Test Score

	(1) Any College	(2) 2yr Public	(3) 4yr Public	(4) Selective	(5) Grad Any	(6) Wage 2022	(7) Income Percentile
Proportion Upper-Income	0.0516 (0.0240)	0.0422 (0.0245)	0.0377 (0.0185)	0.00401 (0.00584)	0.0358 (0.0223)	1057.0 (402.0)	4,609 (1,680)
Proportion High Achieve (G3 Reading)	-0.00340 (0.0210)	-0.000834 (0.0227)	0.0121 (0.0156)	-0.00392 (0.00468)	0.0122 (0.0224)	-699.9 (319.1)	-3,595 (1,417)
N Clusters	2141	2141	2141	2141	1992	1992	1992
N	1013905	1013905	1013905	1013905	700245	700245	700245

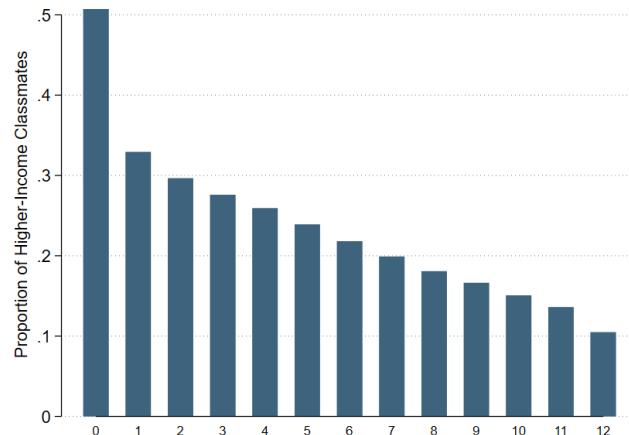
Standard errors are in parentheses and clustered at the school level. The coefficients are based on Equation (2) but defining a student as high-achieving based on grade 3 instead of grade 8 reading test-scores and controlling for students' own grade 3 reading and math test-scores instead of grade 8. Models (1)–(4) include cohorts 2010–2016. Models (5)–(7) include cohorts 2010–2014.

Figure A1: Pattern of Exposure to High-Income Peers is Consistent Independent of Income Measure: Higher-Income Defined as Above 150/200 Thousand



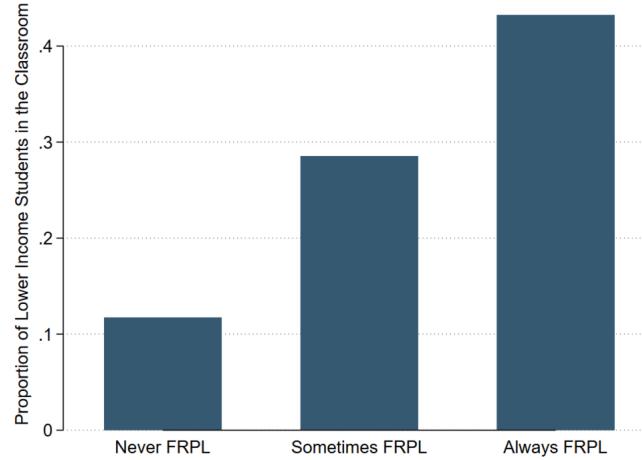
*Notes:* Parental income is based on parental gross adjusted income listed on the financial aid application. There may be students with parental income equivalent to this or higher but who have not applied for financial aid; 4% and 2% of students have financial aid data and parental income above \$150k and above \$200k. Of those who have financial aid data, 11% have parental income above \$150k and 6% have parental income data above \$200k.

Figure A2: Exposure to Higher-Income Peers Decreases with Number of Years on Free/Reduced Lunch



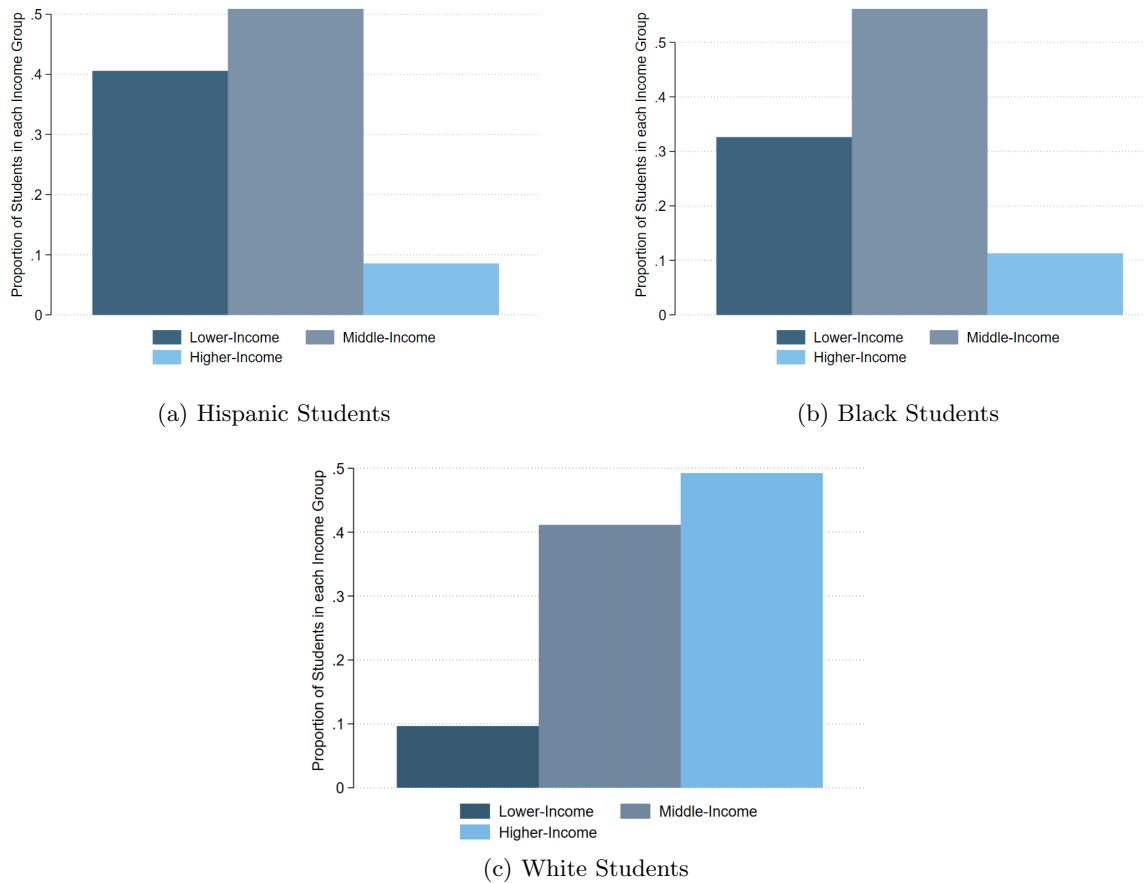
*Notes:* The X-axis captures years on free/reduced lunch. Years on free/reduced lunch depends on students' free/reduced price lunch status in each of the sixteen years I observe them in the data. I do not observe most students for more than thirteen years, and thus, I cap the number of years on free/reduced price lunch status at twelve. Twelve years includes students who are observed for 12 or more years on free/reduced price lunch. Exposure to higher-income peers is the proportion of total classmates between grades five to twelve that are higher income.

Figure A3: Proportion of Lower-Income Students in the Classroom: by Income Group



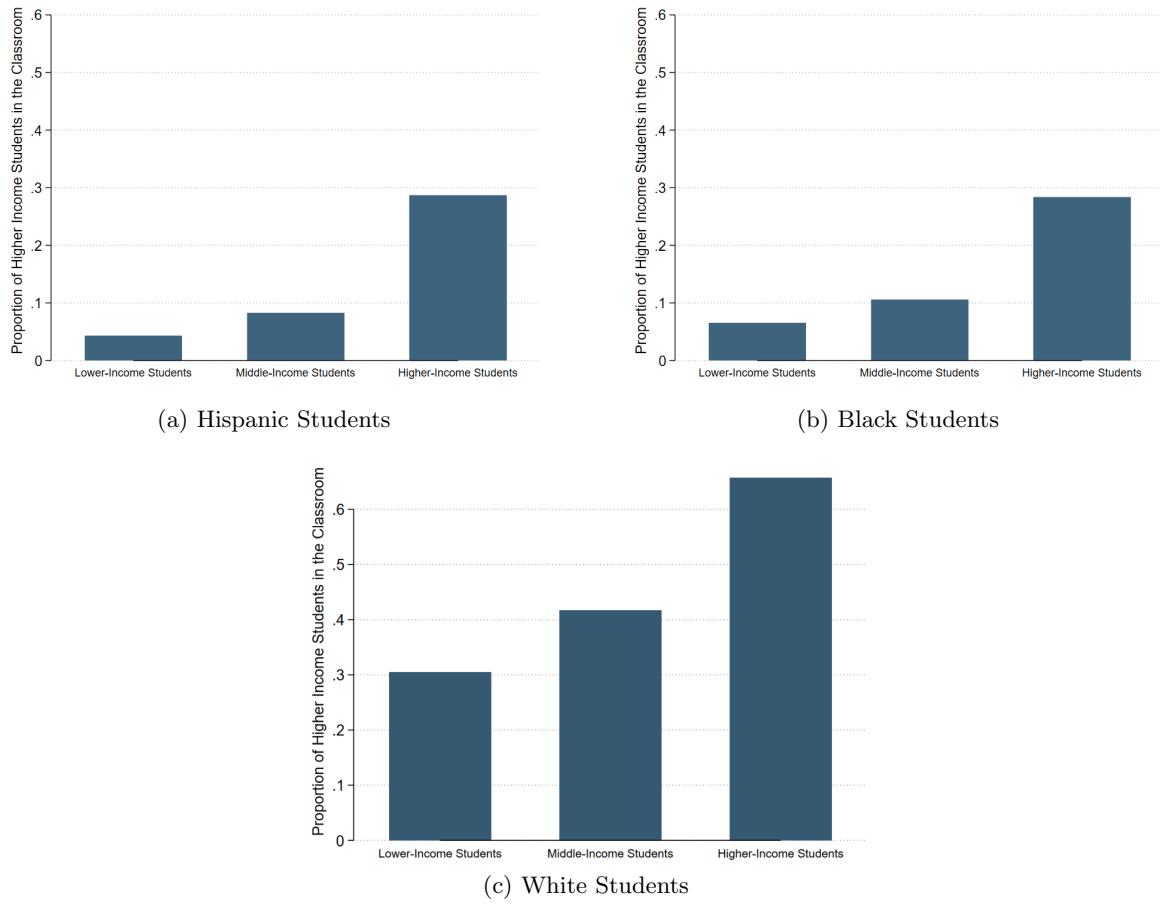
*Notes:* lower-income students are defined as students always on free/reduced lunch.

Figure A4: Distribution of Parental Income by Race/Ethnicity



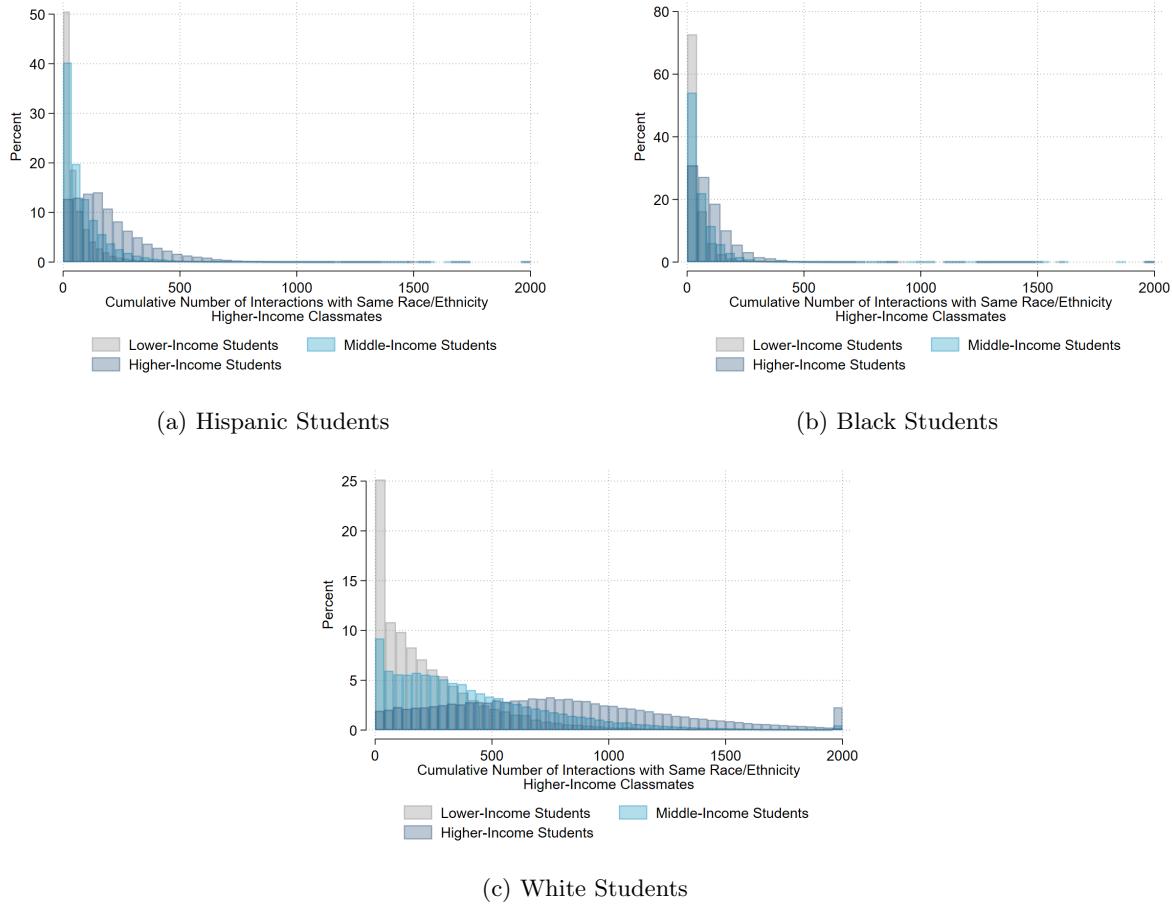
*Notes:* This plot captures the average proportion of students in each income level, for each racial/ethnic group across three cohorts. The average is weighted by the number of students across three cohorts in each racial/ethnic group.

Figure A5: Proportion of Higher-Income, Same-Race/Ethnicity Classmates



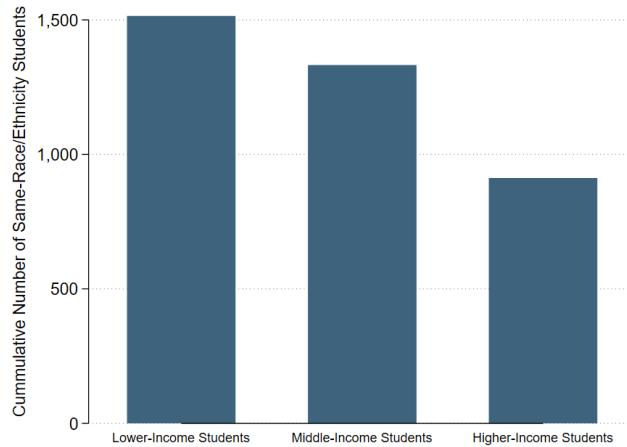
*Notes:* These plots capture the proportion of higher-income classmates of the same race/ethnicity across grades 5 to 12 for each income and racial/ethnic group. The average is weighted by the number of students across three cohorts in each racial/ethnic group.

Figure A6: Cumulative Number of Higher-Income Same-Race/Ethnicity Classmates

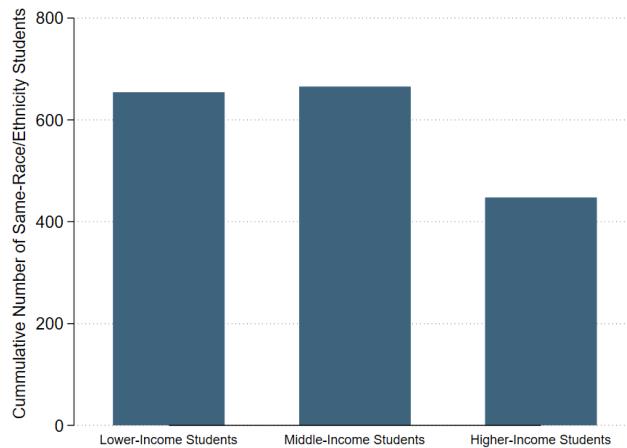


*Notes:* Higher-income students are defined as students never on free/reduced lunch. These plots capture the percentage of students in each income group with various cumulative interactions with higher-income students. Outlier observations with more than 2000 interactions with higher-income classmates of the same race/ethnicity are top-coded. Every time a student is in a classroom with a higher-income student, it is counted as one interaction.

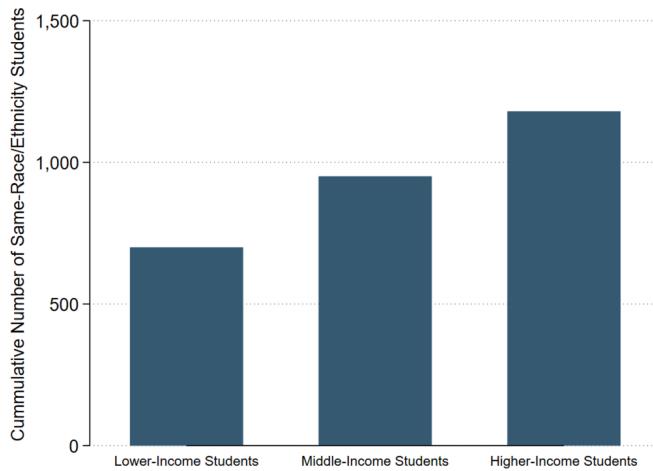
Figure A7: Cumulative Average Number of Same-Race/Ethnicity Classmates



(a) Hispanic Students



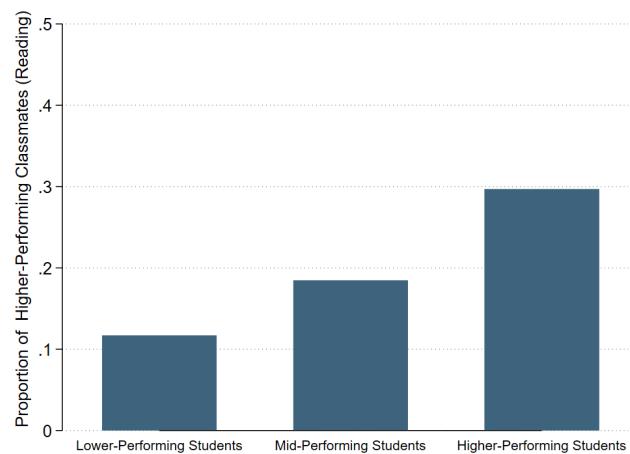
(b) Black Students



(c) White Students

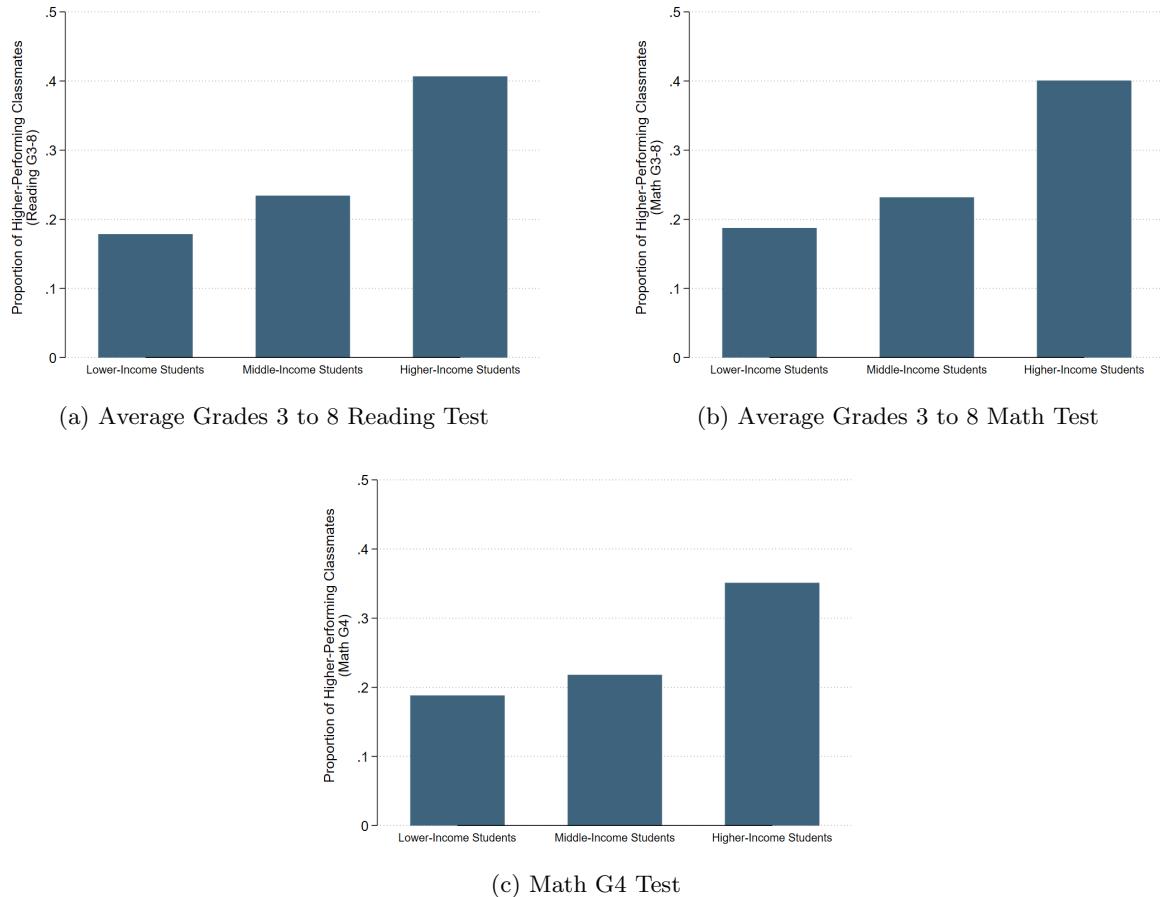
*Notes:* Higher-income students are defined as students never on free/reduced lunch. Each plot presents the average number of students of the same race/ethnicity a student is likely to interact with across grades 5 to 12.

Figure A8: The Cross-Test-Score Exposure Gap is Smaller than the Cross-Income Exposure Gap



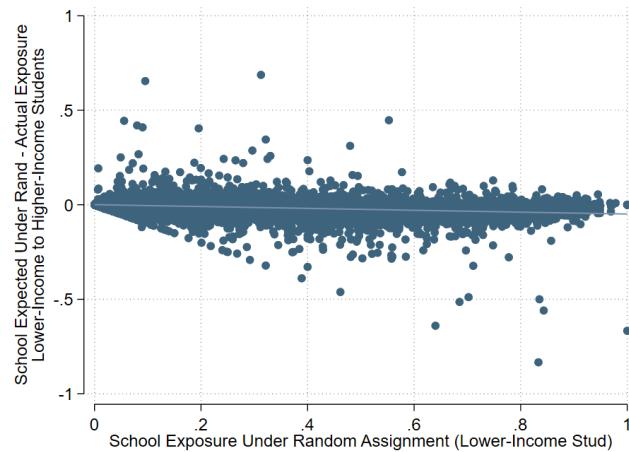
*Notes:* This plot captures students' cumulative proportion of higher-performing classmates between grade 5 and expected grade 12. I define higher-performing students as students who performed in the top 24 percentiles of the test score based on the grade 4 standardized reading test. The test score is standardized based on the raw score for all students who have taken the test in a given year. The percentile is based on the distribution of students who have a grade 4 test score. Approximately 5% of students are missing grade 4 reading test scores. I impute students' missing test score with 0 (i.e., the average test-score because the test score is standardized). Lower-achieving students are those who performed in the bottom 29 percentiles of grade 4 test scores.

Figure A9: Sensitive to Choice of Test: Measure of Exposure to Higher-Achieving Students



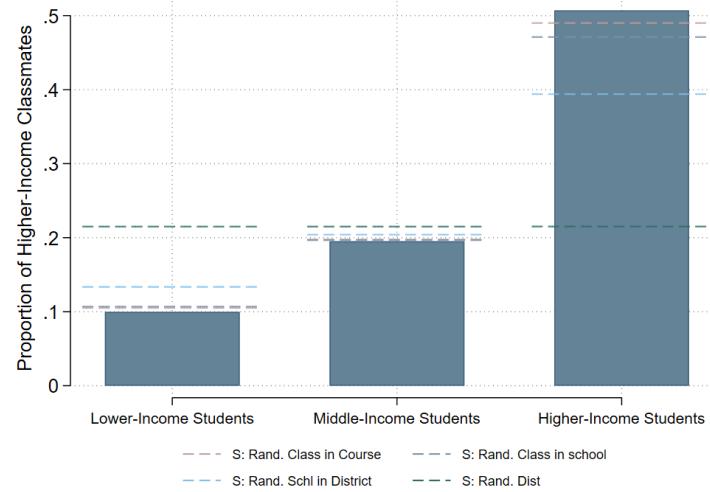
*Notes:* Plots capture the average cumulative proportion of total classmates that are higher-achieving by student income. All tests are standardized within year-grade. High achieving students are those who are in the top 24 percentiles of the average grades 3 to 8 standardized tests (or grade 4 test in panel (c)). Cumulative exposure based on following three cohorts of grade 5 students (2012-2014) to expected grade 12.

Figure A10: Relationship Between the Gap in Exposure and Potential Exposure Under Random Assignment



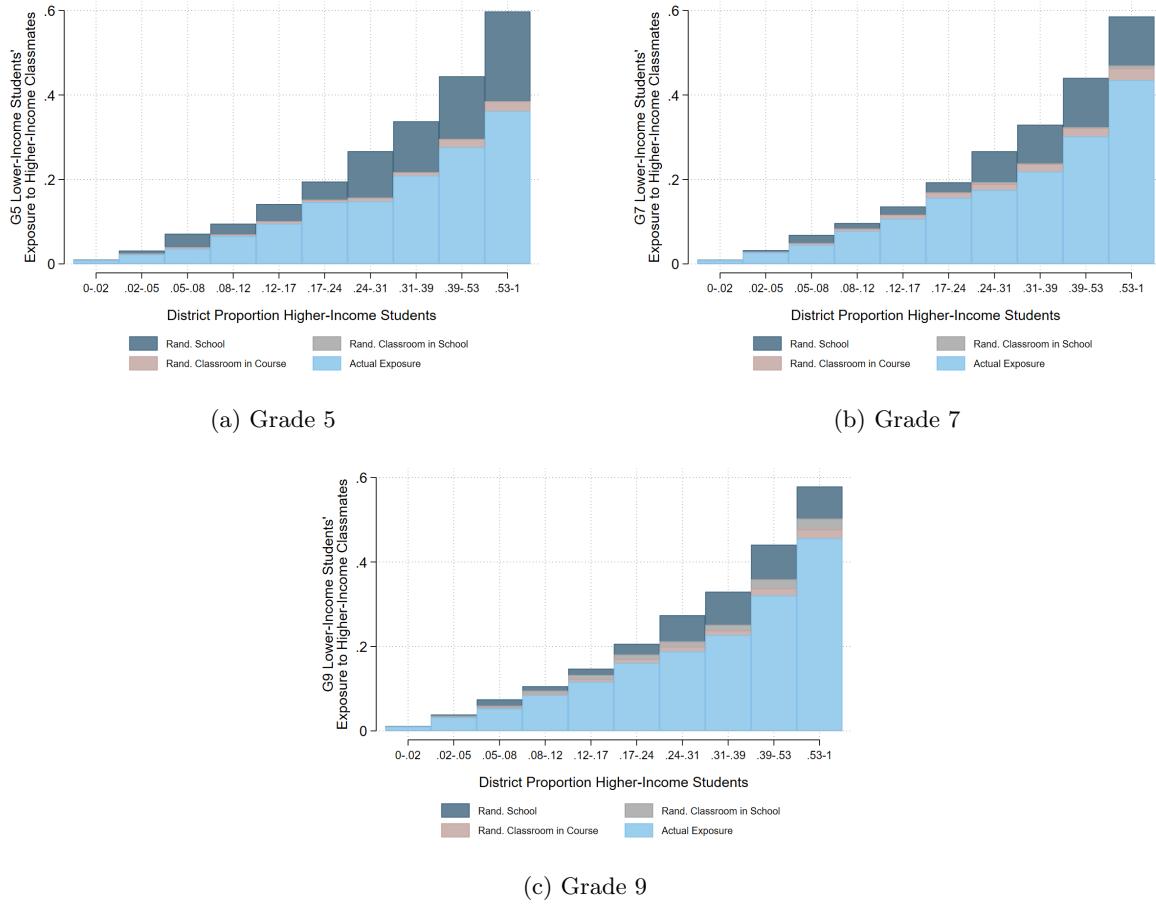
*Notes:* The scatter plot is at the cohort-year-school level. On the x-axis is the potential exposure to higher-income students under random assignment to classrooms within a school. This is on average equal to the proportion of higher-income students in the school. The y-axis is the difference between actual exposure to higher-income students and expected exposure had students been randomly assigned to classrooms. A larger value means that lower-income students are less likely to be exposed to higher-income students than we would have expected under random assignment to classrooms.

Figure A11: Exposure to Higher-Income Students: Observed Relative to Random Assignment



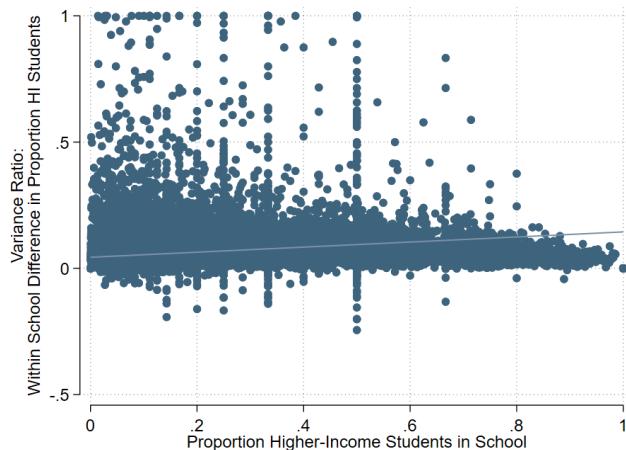
*Notes:* Higher-income students are defined as students never on free/reduced lunch. The bars capture the cumulative proportion of classmates that are higher-income calculated as the total number of higher-income students a given student is in a classroom with from grade 5 to expected grade 12 (excluding own status) divided by the total number of students in each classroom from grade 5 to expected grade 12 (excluding self). The average exposure is weighted by the number of students in each group. The dashed lines capture the various integration benchmarks. The “S: Rand. Dist” lines capture the expected proportion of higher-income classmates had students been randomly assigned to districts in each year-cohort: district integration benchmark. The “S: Rand Schl in District” lines capture the expected proportion of higher-income classmates had students been randomly assigned to schools within a district in each year-cohort (holding constant district composition): school integration benchmark. The “S: Rand Class in School” lines capture the expected proportion of higher-income classmates had students been randomly assigned to classrooms within a school in each year-cohort (holding constant school composition): classroom integration benchmark. The “S: Rand Class in Course” lines capture the expected proportion of higher-income classmates had students been randomly assigned to classrooms within a school in each year-cohort (holding constant school-course composition).

Figure A12: Choice of Course Accounts for More of the Gap in Exposure in Grade 9 than in Earlier Grades



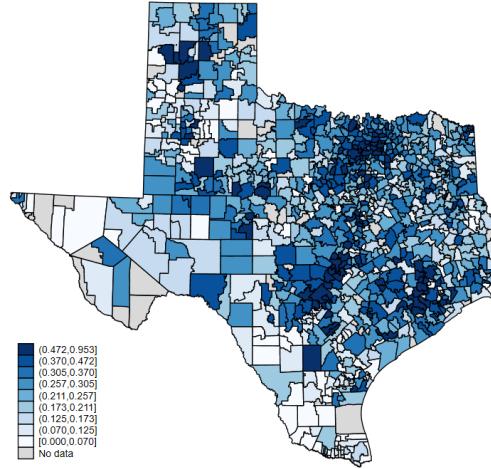
*Notes:* This plot presents lower-income students' exposure to higher-income students in districts with various proportions of higher-income students. Exposure is defined as the average proportion of higher-income classmates in a year. The light blue bar shows the observed proportion of higher-income students in lower-income students' classrooms. The navy bar presents the expected proportion of higher-income students had students been randomly assigned to schools in the district: school integration benchmark. The gray bar presents the expected proportion of higher-income students had students been randomly assigned to classrooms in the school: classroom integration benchmark. The pink bar presents the expected proportion of higher-income students had students been randomly assigned to classrooms in a course in a school. Districts are split into ten percentiles based on the distribution of the proportion of higher-income students in the district in a given school-year. The distribution is weighted by the number of students in each district (independent of income status). The lower and upper bound for each of the percentiles is shown on the x-axis. The panels present the gap by expected grade.

Figure A13: Correlation between the School Variance Ratio and Proportion of Higher Income Students in the School

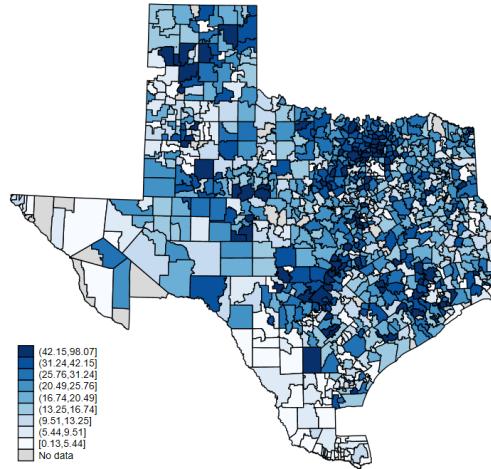


*Notes:* The y-axis captures the difference in the proportion of higher-income students in higher- relative to other-income student classrooms. The x-axis presents the proportion of higher-income students in the school across cohorts.

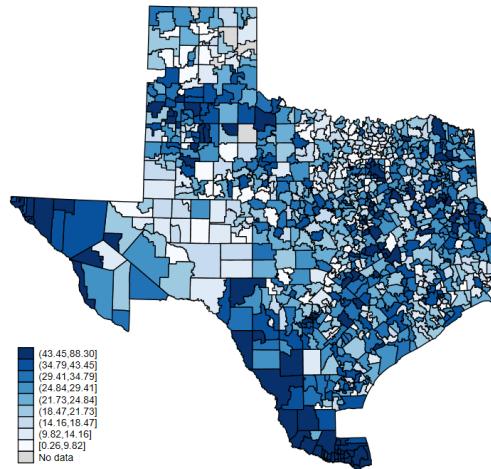
Figure A14: Percentage of Higher- and Lower-Income students



(a) Percentage of LI Students' Classmates that are Higher-Income



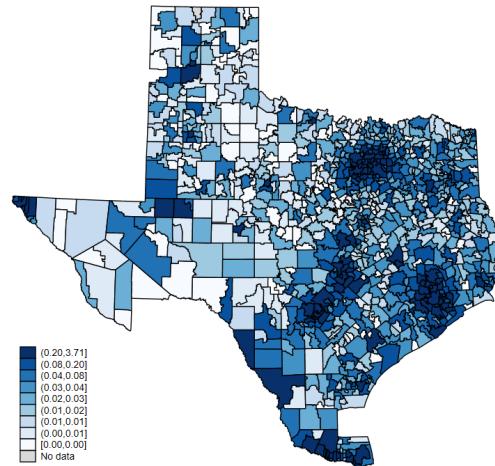
(b) Percentage of Students that are Higher-Income



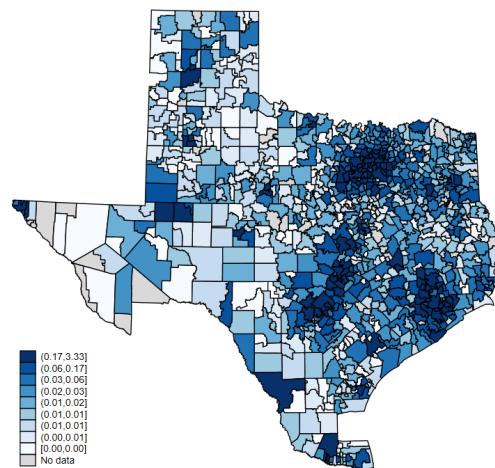
(c) Percentage of Students that are Lower-Income

*Notes:* The map in Panel (a) captures the percent of classmates in lower-income students' classrooms. The map in Panel (b) captures the percentage of students in each district that are higher income. Panel (c) captures the percentage of students that are lower income. In Panels (b) and (c), areas with fewer than 10 students in the numerator are shaded in grey under the "No-data" category. In Panel (a), areas with fewer than 10 higher- or lower-income students are shaded in grey under the "No-data" category.

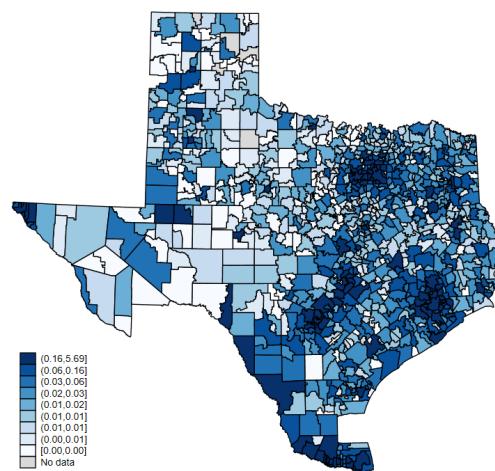
Figure A15: Percentage of Population in Each District



(a) Percentage of Total Student Population



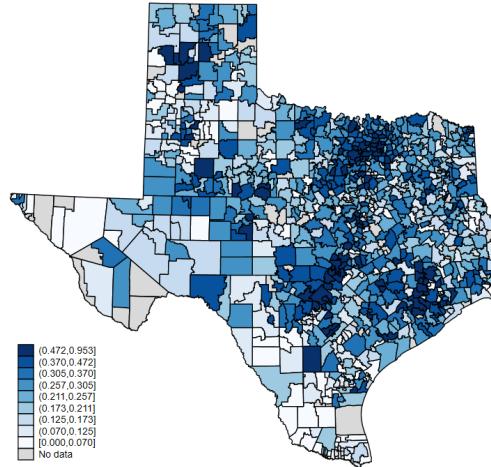
(b) Percentage of Total Higher-Income student Population



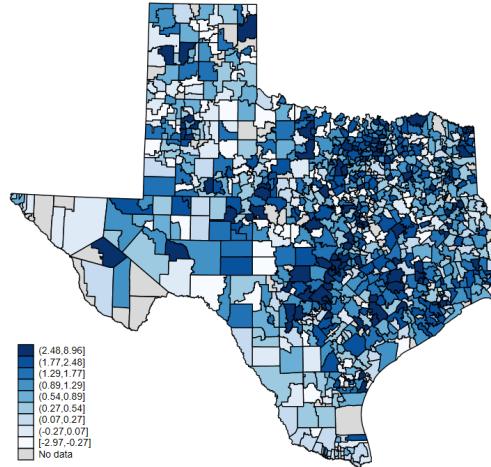
(c) Percentage of Total Lower-Income student Population

*Notes:* Panel (a) captures the percentage of the total student population residing in each district. Panel (b) captures the percentage of the total higher-income student population residing in each district. Panel (c) captures the percentage of the total lower-income student population residing in each district. Areas with fewer than 10 students are shaded in grey under the “No-data” category.

Figure A16: Observed Exposure Relative to Exposure Under Random Assignment



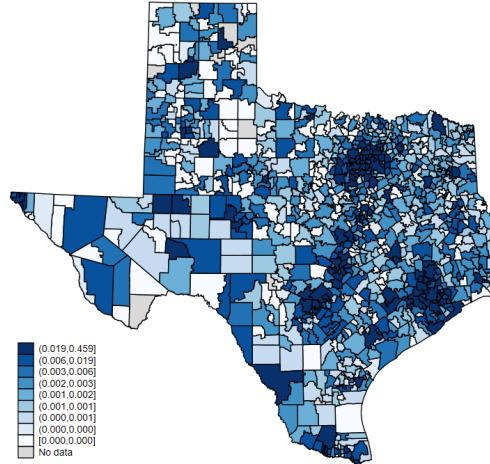
(a) Relative to Random Assignment to Schools and Classrooms within Districts



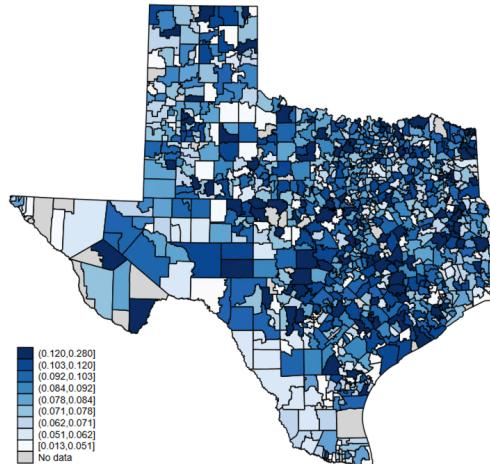
(b) Relative to Random Assignment to Classrooms within Schools

*Notes:* This plot is limited to lower-income students. Lower-income students are students always on free/reduced lunch. The units used are percentages. The plot captures the percentage-point difference between the observed proportion of higher-income students to whom lower-income students are exposed and the expected proportion of higher-income students had students been randomly assigned to classrooms within the district. The maps are constructed by grouping school districts into 9 deciles and shading the areas so that lighter colors correspond to lower difference in observed and simulated exposure. Areas with fewer than 10 higher- or lower-income students are shaded in grey under the “No-data” category. Figure (a) captures the percentage-point difference between the observed proportion of higher-income students to whom lower-income students are exposed and the expected proportion of higher-income students had students been randomly assigned to schools and classrooms within the district. Figure (b) captures the percentage point difference between the observed proportion of higher-income students to whom lower-income students are exposed and the expected proportion of higher-income students had students been randomly assigned to classrooms within the district.

Figure A17: Level of Sorting by Income Within and Between Schools in a District



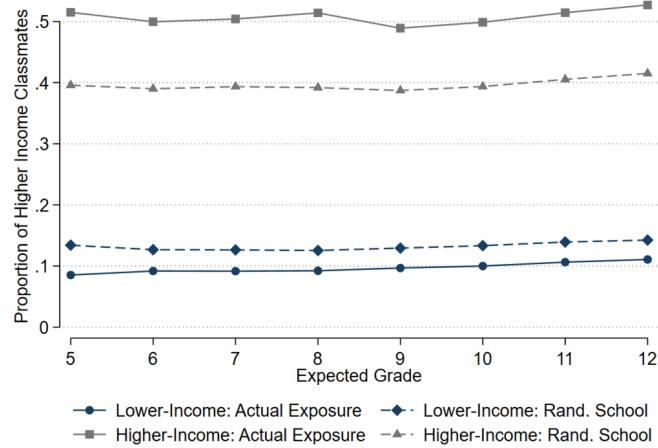
(a) Variance Ratio between Schools



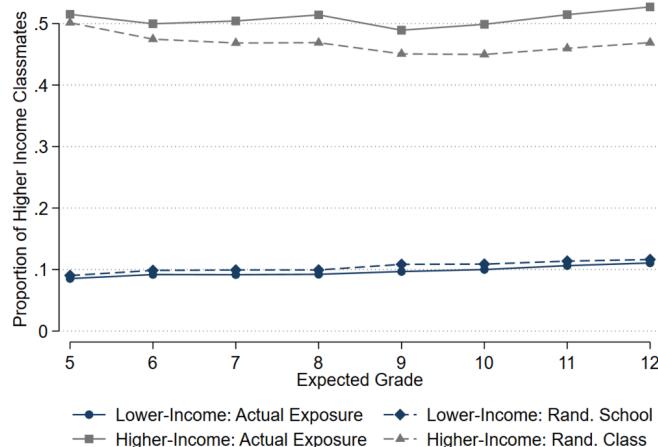
(b) Variance Ratio within Schools

*Notes:* Panel (a) captures within-district sorting across schools by income. It is the difference in the proportion of higher-income students in higher- relative to other-income student schools, within the same district. Panel (b) captures within-school sorting across classrooms. The line of best fit is weighted by the number of students enrolled across cohorts in each expected grade level. It is the difference in the proportion of higher-income students in higher- relative to other-income student classrooms, within the same school.

Figure A18: The Exposure Gap between Random School Assignment and Observed is Consistent Across Grades but Increases under Random Classroom Assignment



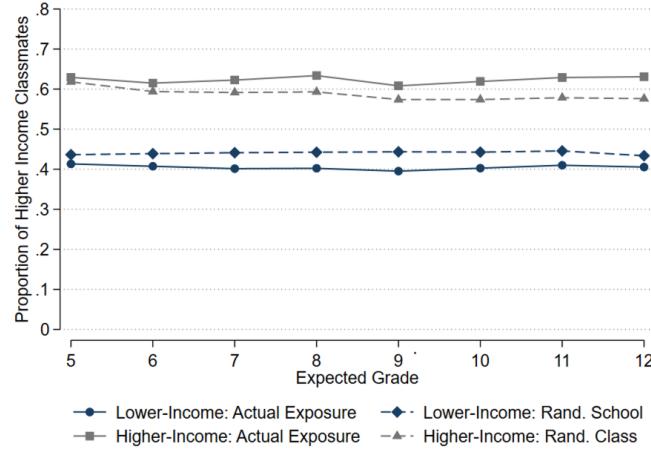
(a) Random Assignment to Schools and Classrooms within District



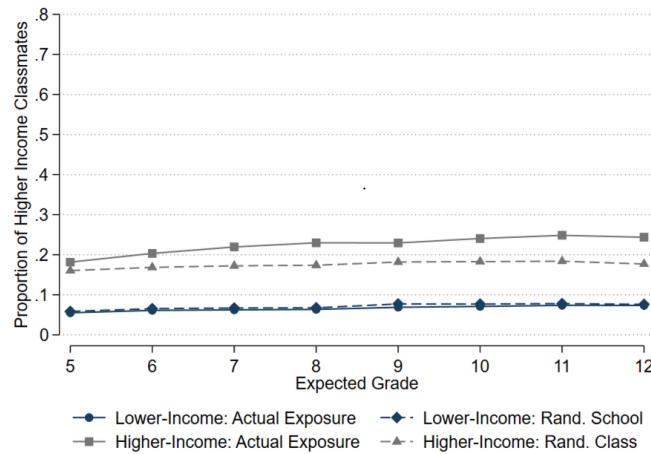
(b) Random Assignment to Classrooms within School

*Notes:* Higher-income students are defined as students never on free/reduced lunch. The solid lines in both plots capture exposure to higher-income students in each grade. Exposure is defined as the proportion of a student's classmates that are higher income. Expected grade is based on what year it is and when I first observe a cohort of students in grade 5. The average exposure is based on the number of students across cohorts enrolled in an expected grade. In Panel (a), the dashed lines capture the average proportion of higher-income students had students been randomly assigned to schools and classrooms in a district each year. In Panel (b), the dashed lines capture the average proportion of higher-income students had students been randomly assigned to classrooms in a school each year.

Figure A19: Classroom Sorting Matters for Lower-Income Students in Schools with a Large Proportion of Higher-Income Students



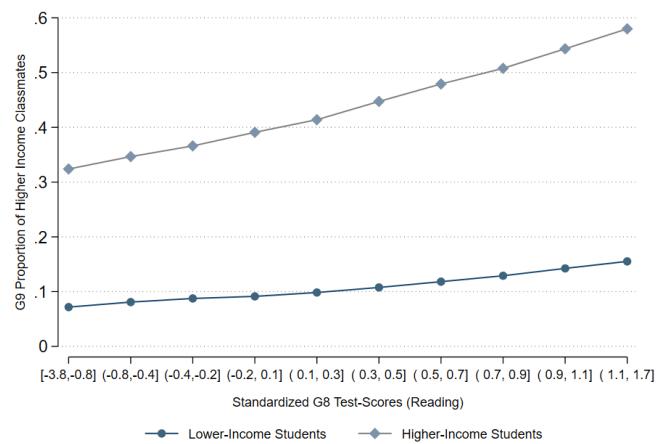
(a) Higher-Income Schools



(b) Lower-Income Schools

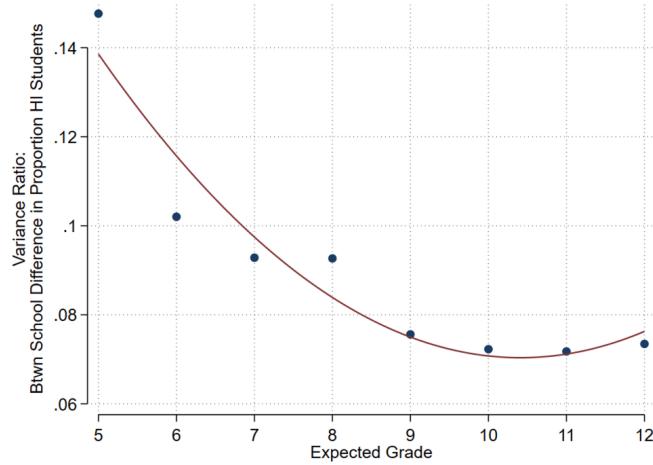
*Notes:* Higher-Income schools are schools that are in the top 70th percentile of schools in terms of the proportion of higher-income students they serve. Those are schools that serve 30% or more higher-income students. The dashed lines in both plots capture the average proportion of higher-income students had students been randomly assigned to classrooms in a school each year. The solid lines in both plots capture exposure to higher-income students in each grade. Exposure is defined as the proportion of a student's classmates that are higher income. Expected grade is based on what year it is and when I first observe a cohort of students in grade 5. The average exposure is based on the number of students across cohorts enrolled in a expected grade.

Figure A20: Gap in Exposure to Higher-Income Students Exists Independent of Academic Performance

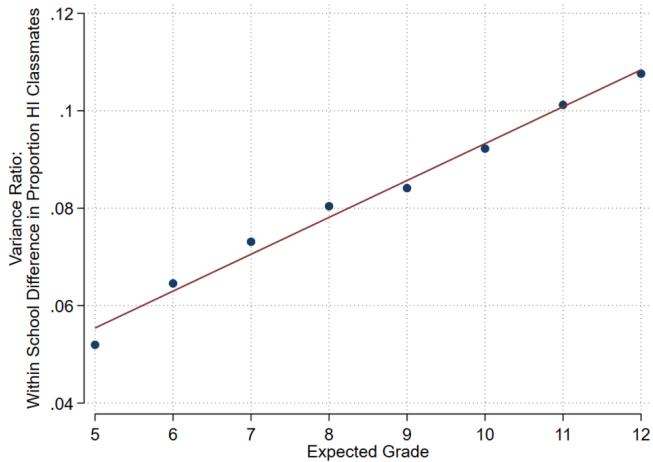


*Notes:* The y-axis captures students' average exposure to higher-income classmates (excluding own status) in expected grade 9 in each bin. Student test score is based on the grade 8 reading test, standardized based on the full sample of students who have taken the test in a given year. The plot excludes students without test scores (7% of the sample of students are missing grade 8 reading test scores). Students are split into 10 percentiles based on their test score in grade 8. The range of standardized test scores included in each percentile is presented on the x-axis.

Figure A21: Sorting between Schools Decreases as Within-School Sorting Increases Linearly with Every Grade



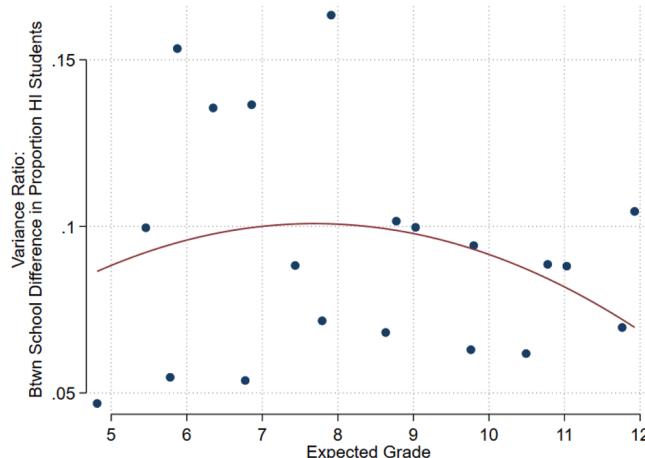
(a) Within District, Difference in School Proportion of Higher-Income Students



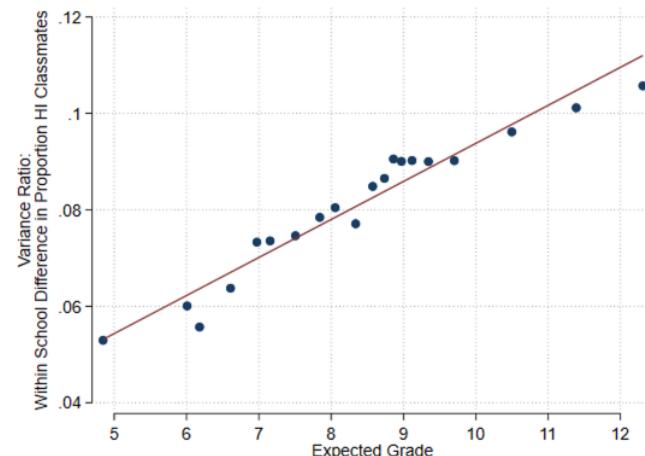
(b) Within School, Difference in Classroom Proportion of Higher-Income Students

*Notes:* Panel (a) captures the relationship between within-district sorting across schools and grade level. Panel (b) captures the relationship between within-school sorting across classrooms and grade level. The line of best fit is weighted by the number of students enrolled across cohorts in each expected grade level. Bins are based on grouping the x-values into 20 equal-sized bins. It then computes the average y-variable value for each bin.

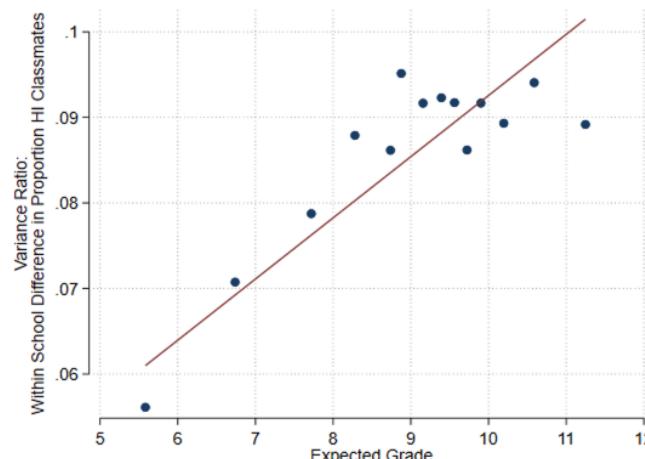
Figure A22: Change in Sorting between Schools is Mainly Driven by the Change in the Number of School Options



(a) Within District, Difference in Proportion of HI Students in School: Conditional on School-to-Students



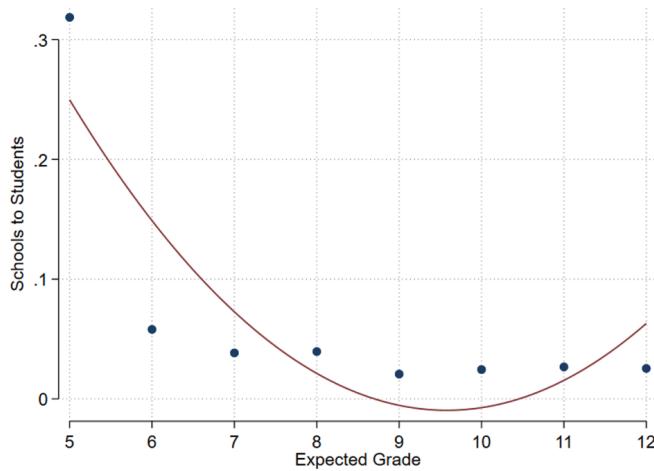
(b) Within School, Difference in Proportion of HI Students in Classroom: Conditional on Classroom-to-Students



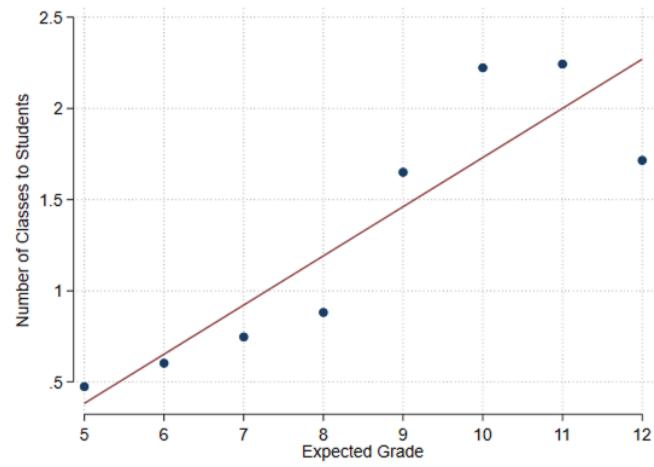
(c) Within School, Difference in Proportion of HI Students in Classroom: Conditional on Advanced Courses-to-Students

*Notes:* Panel (a) captures the relationship between within-district sorting across schools and grade level, with controls for the number of schools serving that grade level in a district to the number of students enrolled in that grade level. Panels (b) and (c) capture the relationship between within-school sorting across classrooms and grade level. In Panel (b), I control for the number of classrooms to students offered in a school at a grade level. In Panel (c), I control for the number of advanced courses offered in a grade level in a school. Note that courses are labeled as advanced only at the high school level. The line of best fit is weighted by the number of students enrolled across cohorts in each expected grade level. Bins are based on grouping the x-values into 20 equal-sized bins. It then computes the average y-variable value for each bin.

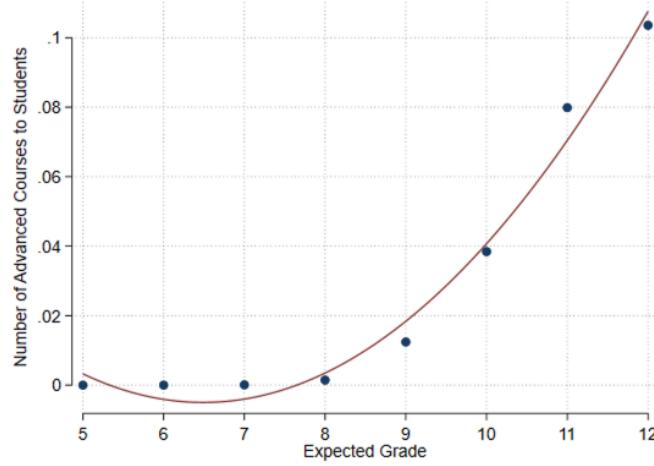
Figure A23: Changes in School- and Class-to-Student Ratio Across Grades



(a) School-to-Student Ratio Across Grades



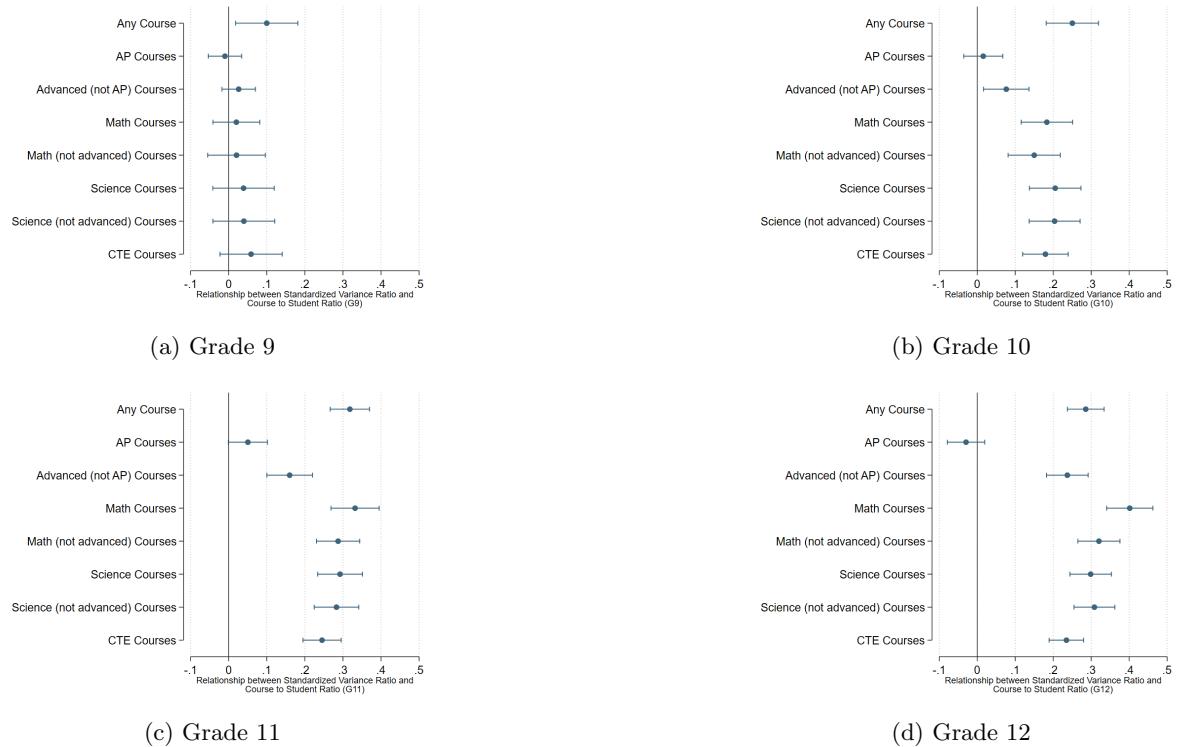
(b) Classroom-to-Student Ratio Across Grades



(c) Advanced Courses-to-Student Ratio Across Grades

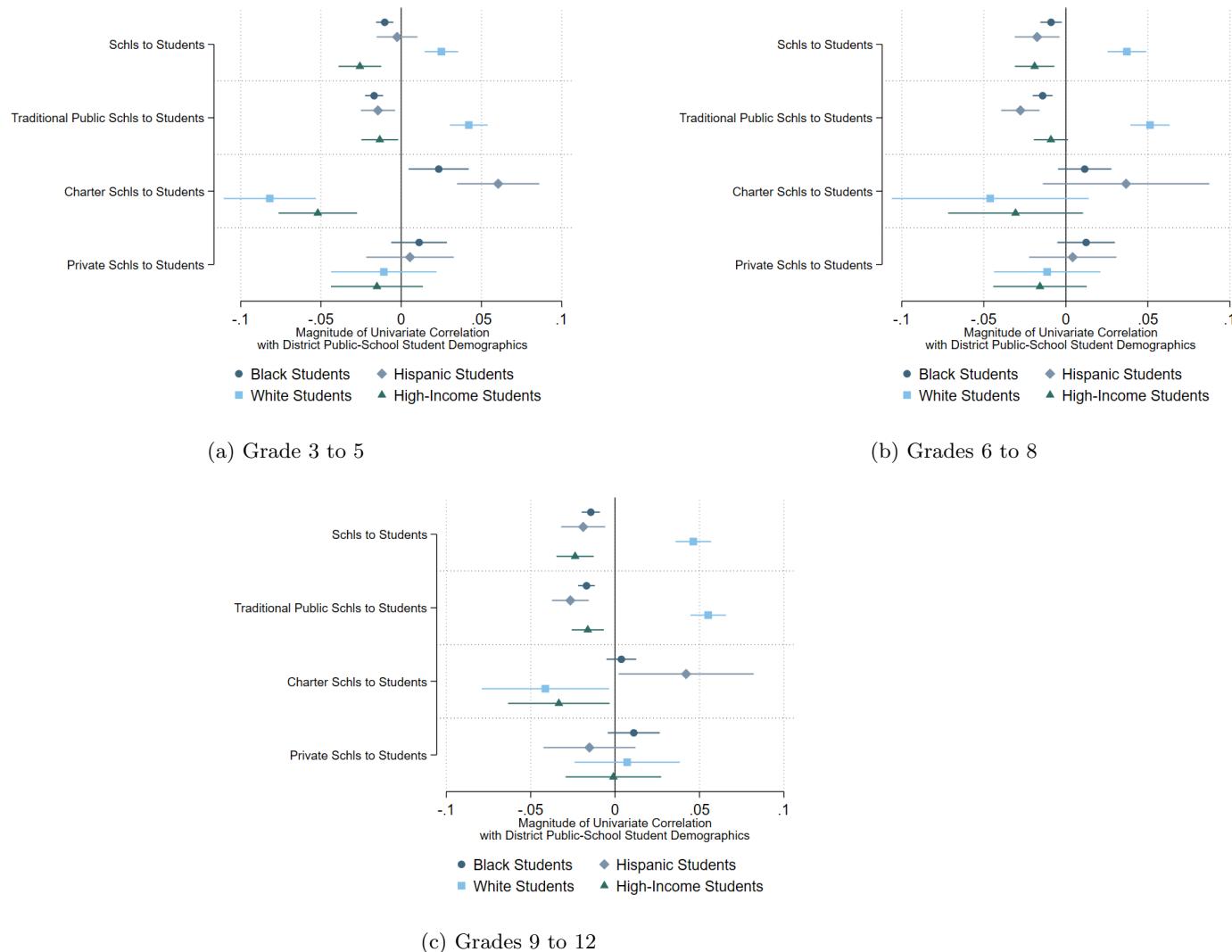
Notes: Panel (a) captures the average number of schools to students in each grade level. Panel (b) captures the average number of classrooms to students in each grade level. Panel (c) captures the average number of advanced courses to students in each grade level. Note that courses are identified as advanced in the TEA data starting in high school.

Figure A24: Type and Number of Courses Offered Seem to Account for More of the Variation in Sorting in Later High School Grades



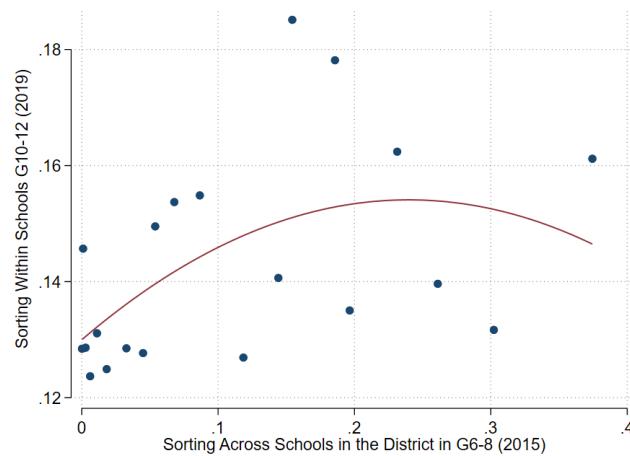
*Notes:* Based on 2019 high school classroom enrollment data. School-level univariate regression coefficients from regressing the standardized sorting measure on the number of courses offered to students served are presented standardized in the first row. The other rows capture the correlation with other specific groups of courses to students served. Math courses capture the number of any math course to number of students served, standardized. Math (not advanced) courses capture all math courses that are not advanced to the number of students served. The standardization is based on the weighted distribution of courses to students and level of within-school sorting in each grade across schools. The correlation is weighted by the number of students enrolled in a school in a given grade. The 95% confidence intervals are presented. The standard errors are clustered at the school level.

Figure A25: District Demographics and School Offerings



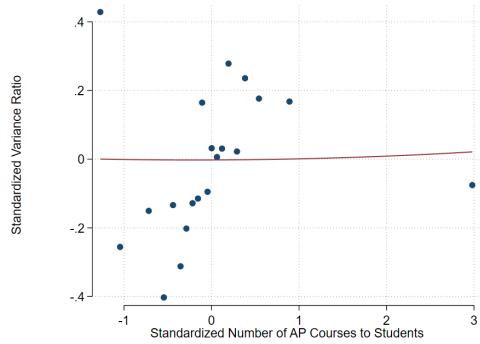
*Notes:* The number of schools to students is standardized – a one-unit increase is a one-standard-deviation increase in the number of schools to students. Schools are placed into districts, and the ratio is calculated as the number of schools in the district to the number of students served in the district in 2019. The number of private schools and students served in these schools is estimated based on a list of private schools and the number of students and grades served by each school obtained from the Texas Alliances Accredited Private Commission's archive for the 2018–2019 school year. The estimates are based on regressing the proportion of a given demographic (e.g., proportion of Hispanic students) on the standardized number of schools to students in the county.

Figure A26: The Level of Between-School Sorting in Middle School Does not Seem to Matter for Within-School Sorting in High School

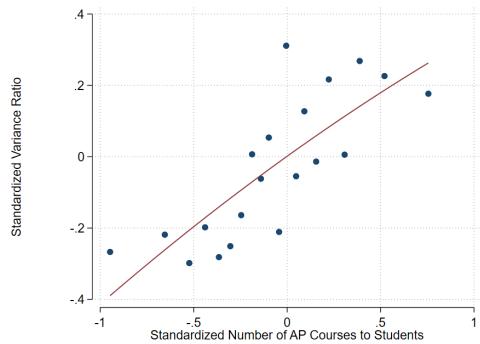


*Notes:* The x-axis captures between-school sorting in middle school in a given district. I use the variance ratio to capture between-school sorting in middle school in 2015. To calculate the between-school variance ratio, I calculate the proportion of higher-income students in the district and the proportion of higher-income students in middle schools in which higher-income students are enrolled. The variance ratio captures the average difference in the proportion of higher-income students in schools attended by higher- relative to other-income students. The y-axis captures within-school difference in the proportion of higher-income students in higher- relative to other-income student classrooms as captured by Equation (1). Of students in middle school in 2015, 78% are enrolled in the same district in 2019. The fitted line is based on weighting each school by the number of students enrolled in a school in high school. Bins are equal sizes, based on the number of observations divided by 20. Because charter schools are not assigned to a district, this plot excludes students enrolled in charter schools. This is a small number of students (approximately 5% of students).

Figure A27: Relationship between AP Courses to Students and Variance Ratio



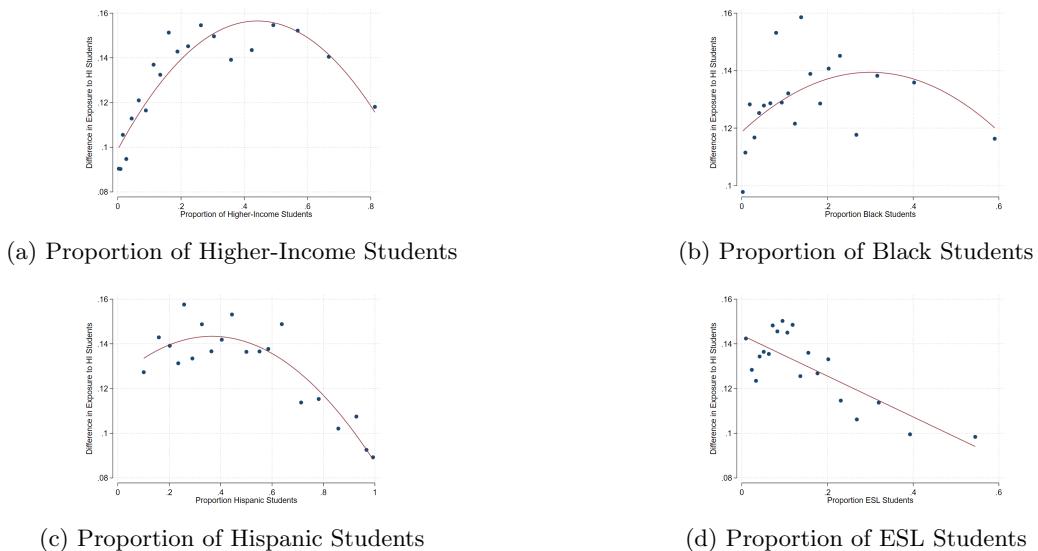
(a) Full Sample



(b) Excluding above 1 standard deviation AP courses to students and schools with no AP course offered

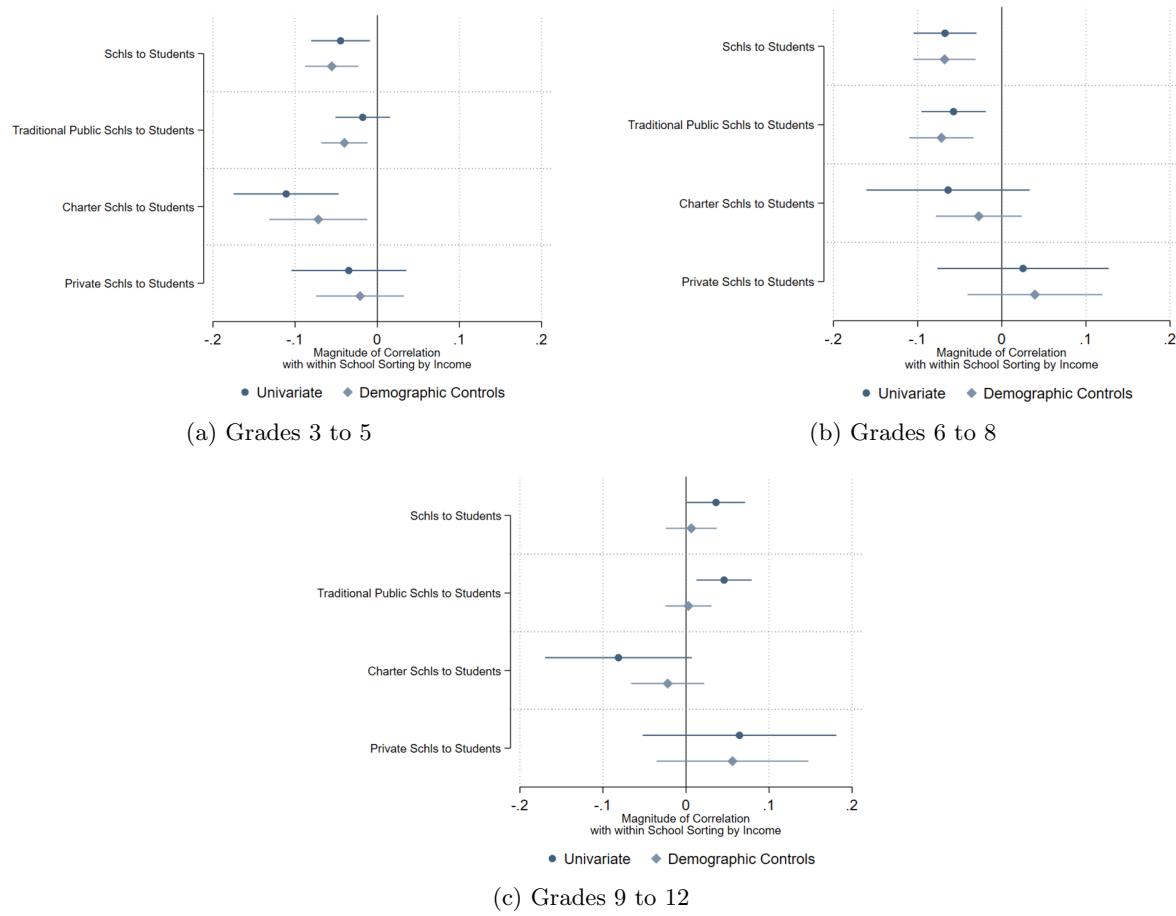
*Notes:* Based on 2019 high school classroom enrollment data. Bins are based on grouping the x-values into 20 equal-sized bins. It then computes the average y-variable value for that bin. The fitted line is weighted by the number of high school students enrolled. Panel (a) plots the relationship between the number of AP courses to students and the full sample. Panel (b) limits the sample to schools offering at least one AP course and below 1 standard deviations in terms of the number of AP courses they offer to students.

Figure A28: School Composition and Within-School Sorting by Income



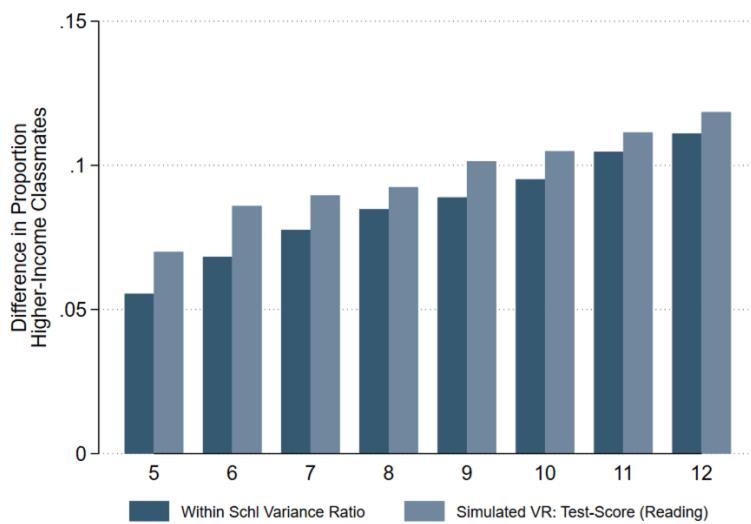
*Notes:* Based on 2019 high school classroom enrollment data. Panel (a) captures the relationship between the proportion of higher-income students and the within-school sorting. Panel (b) captures the relationship between the proportion of Black students and sorting. Panel (c) captures the relationship between the proportion of Latin-American students and sorting. Panel (d) captures the relationship between the proportion of students with an English as a Second Language designation and within-school sorting. Bins are based on grouping the x-values into 20 equal-sized bins. It then computes the average y-variable value for that bin. The fitted line is weighted by the number of high school students enrolled.

Figure A29: No Strong Evidence of an Association between Within-School Sorting in Public Schools and Number of School Options



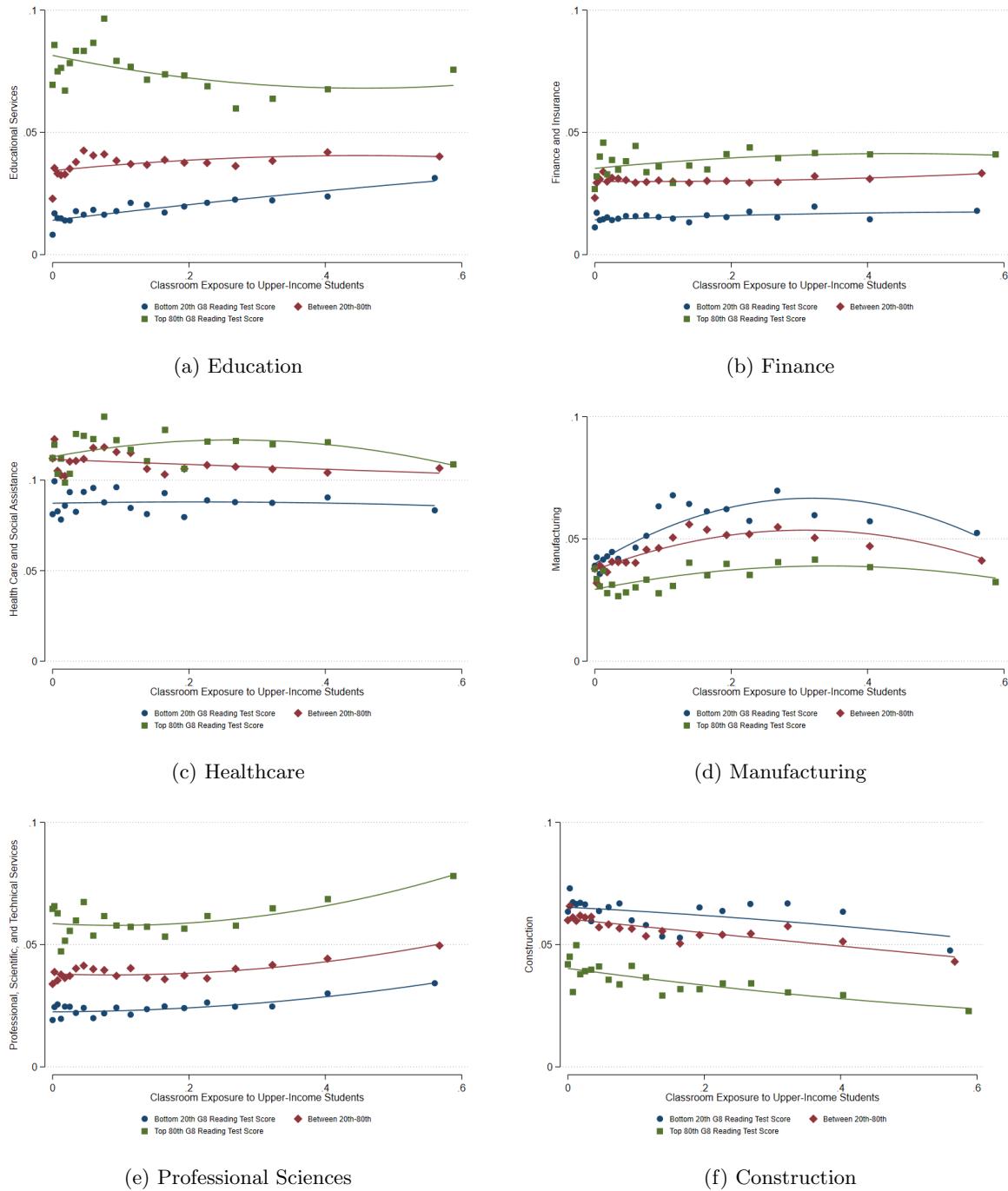
*Notes:* Similar to Figure 9 but with the within-school variance ratio (sorting) as the outcome.

Figure A30: Had Students Been Assigned to Classrooms by Test Score, the Gap in Exposure by Income Within Schools Would Have Been Slightly Larger



*Notes:* This plot captures the variance ratio (VR) under two settings. The Within Schol Variance Ratio bars capture the observed within-school variance ratio in every grade. The Simulated VR: Test Score bars capture the variance ratio under a simulation in which students in a school are ranked by grade 4 test score and then placed in classrooms by test score. I keep the number of classrooms and students assigned to a classroom the same as that observed in a given year. I run the simulation using students' grade 4 reading test scores. School grade levels that either do not have higher-income students or do not have other-income students are not included in the sample to focus on the role of test scores in schools with variation in income composition.

Figure A31: Exposure to Upper-Income Students and Industry of Employment



*Notes:* The line and plots are similar to Figure (13). All outcomes include cohorts 2011 to 2014. The industry of employment is based on unemployment insurance data and does not identify the type of job within the industry.